KAPOSI'S SARCOMA-ASSOCIATED HERPESVIRUS REPLICATION AND TRANSCRIPTION ACTIVATOR REGULATES EXTRACELLULAR MATRIX SIGNAL PATHWAY

By

Daniel Pfalmer

RECOMMENDED:

Dr. Karsten Hueffer
Advisory Committee Member

Dr. Andrea Ferrante
Advisory Committee Member

Dr. Jiguo "Jack" Chen
Advisory Committee Chair

Dr. Diane Wagner
Chair, Department of Biology and Wildlife

APPROVED:

Dr. Paul Layer
Dean, College of Natural Science and Mathematics

Dr. John Eichelberger
Dean of the Graduate School

18 August 2016
Date

KAPOSI'S SARCOMA-ASSOCIATED HERPESVIRUS REPLICATION AND TRANSCRIPTION ACTIVATOR REGULATES EXTRACELLULAR MATRIX SIGNAL PATHWAY

A

THESIS

Presented to the Faculty

of the University of Alaska Fairbanks

in Partial Fulfillment of the Requirements

for the Degree of

MASTER OF SCIENCE

By

Daniel Pfalmer, B.S.

Fairbanks, AK

August 2016

Abstract

Kaposi's Sarcoma (KS) is a malignancy caused by infection with Kaposi's Sarcoma-associated Herpesvirus [KSHV; also known as Human Herpesvirus 8 (HHV8)] in which tumor cells show a characteristic 'spindle-like' morphology. The transcription factor RTA (Replication and Transcription Activator) is the viral protein responsible for reactivating KSHV from its latent state. Production of RTA in latently infected cells causes a number of viral proteins to be produced and leads to a cascade of gene expression changes in both viral and host genes. Previous work in our lab showed that RTA was capable of reprogramming cells *in vitro* to display a spindle-like morphology. In this study we aimed to identify the host gene expression changes caused directly by RTA which could be responsible for that reprogramming. To that end, Madin-Darby Canine Kidney cells (MDCK cells) were chosen as a model for KSHV-naïve mammalian cells. Differences in host gene expression levels in a culture of MDCK cells transfected with a plasmid coding for expression of RTA compared to MDCK cells transfected with a similar plasmid lacking the RTA gene were measured by whole transcriptome sequencing (RNA-Seq). Cells containing the RTA-coding plasmid adopted a spindle-like morphology and showed at least a two-fold change in expression level in approximately 180 genes. Those 180 genes were then screened for known associations to signaling pathways in order to determine which might be involved with the morphological changes observed and/or biological significance. The expression levels of the 10 genes identified by that screening were then verified by quantitative real time PCR (qPCR). Of those 10 genes, eight were identified as potentially associated with the morphological changes, including three genes associated with extra cellular matrix (ECM) destruction (MMP9, CTSD, and CTSS) that were down-regulated; two genes associated with blocking ECM destruction (TIMP1 and TIMP2) that were up-regulated; two ECM component genes (LAMC2 and COL1A2) that were upregulated; and one gene associated with blocking cell-cell and cell-ECM adhesion (MUC1) that was downregulated. The remaining two genes (MAP2K1 and podoplanin) were identified as potentially biologically significant, but not directly involved in regulating morphology. MAP2K1 is associated with epithelial dedifferentiation and was down-regulated; and the lymphatic endothelial specific marker podoplanin (PDPN) was up-regulated. Taken together, the differences in morphology

and gene expression between RTA-producing cells and controls suggest a possible role for RTA in the formation of the spindle cells that characterize Kaposi's sarcoma.

Table of Contents

Page

List of Figures

Page

List of Tables

Page

List of Appendices

Page

Chapter 1. General Introduction

Kaposi's sarcoma (KS) is a malignant neoplasm first described in 1872 by Dr. Moritz Kaposi. KS is characterized as a spindle-cell tumor with variable clinical features ranging from minimal mucocutaneous disease to extensive organ involvement. KS can be primarily categorized into 4 types: Classic, Endemic (African), iatrogenically immunocompromised, and AIDS-related. Classic KS was a rare condition affecting elderly men of Mediterranean descent. It presented as highly vascular lesions on the toes and/or soles of the feet. The lesions progressed very slowly, and were rarely fatal. In the 1960's endemic KS was discovered in central Africa [Hutt and Burkitt, 1965]. The African variant was much more malignant, and occurred in much younger people. Most African patients were males under 30 years old. While the disease was common in certain hotspots in Africa, it was extremely rare in the North America until the early 1980s [Beral et al., 1990]. In 1981 groups of young homosexual men were diagnosed with a highly aggressive form of KS in New York and California. This extremely unusual grouping, along with cases of *Pneumocystis carinii* pneumonia infection in other groups of young homosexual men that had previously been healthy, signaled the start of the AIDS epidemic [Curran and Jaffe, 2011]. KS incidence in the United States went from 1-2 per million people per year before the AIDS epidemic, to 47 cases per million people per year afterward [Eltom et al., 2002]. Before effective treatments for HIV were invented, people with AIDS had roughly a 50% chance of developing KS. It became so common among AIDS patients that KS was categorized as an 'AIDS-defining illness', and developing KS was often how people discovered that they had AIDS. In addition to AIDS patients, there was one more group that had a massively higher risk of developing KS: organ transplant recipients taking immunosuppressive drugs to prevent rejection had a 1 in 200 chance of developing KS [Farge, 1993].

The fact that KS risk was increased not just in patients with HIV, but in other immunocompromised groups as well, suggested that it was not HIV itself causing the tumors, but rather the immune suppression from HIV was somehow allowing the tumors to develop. In 1994 a team at Columbia University discovered the virus that was causing KS , and named it Kaposi's Sarcoma-associated Herpesvirus [KSHV; also known as Human Herpesvirus 8 (HHV8)] [Chang et al., 1994]. KSHV is a gammaherpesvirus, related to another well-known tumor virus,

Epstein-Barr virus. In addition to KS, KSHV also causes primary effusion lymphoma (PEL, a cancer of the lymphocytes) also called body cavity-based lymphoma (BCBL) and some forms of severe lymph node enlargement, called Castleman's disease [Cesarman et al., 1995]. PEL is a highly aggressive B-cell tumor that (like KS) appears in severely immunocompromised individuals infected with KSHV. Despite sharing a common cause, the pathologies, morphologies, and prognoses of the two malignancies are very different. KS tumors are made up of characteristic 'spindle cells' which have an endothelial morphology, as well as immature blood vessels, and inflammatory infiltrate [Ensoli and Stürzl, 1998]. KS forms multiple lesions on the skin and oral mucosa, and in late stages can involve the lymph nodes and visceral organs as well. Conversely, PEL does not form solid tumors at all. Instead, it takes the form of round lymphoblastic cells rapidly growing in suspension. While patients diagnosed with classic KS can live for years and will typically die of unrelated causes, PEL patients often die within 6-months of diagnosis, and rarely survive more than a year [Patel and Xiao, 2013].

Another major difference between PEL and KS involves the life cycle of the virus itself. PEL cells all contain intact copies of the viral genome, and can be successfully cultured in the lab to maintain a continuous cell line. In KS tissue, many of the spindle cells do not contain KSHV genome *in situ*. The proportion of spindle cells containing the KSHV genome increases as the tumor progresses. As few as 10% of spindle cells in early stage tumors can be detected with KSHV; but in the later stages that number can be as high as 90% or more [Cancian et al., 2013]. Numerous attempts have been made to generate stable KSHV-infected cell lines from KS spindle cells, but all have failed [Grundhoff and Ganem, 2004]. It was found that the spindle cells could not grow in culture without the addition of exogenous growth factors, and that even with those growth factors the spindle cells rapidly lost the KSHV genome, showing that the infection was not stable (at least *in vitro*) [Grundhoff and Ganem, 2004].

Finally, as far as morphology and cell surface marker proteins are concerned, PEL cells resemble lymphatic B-cells; while KS spindle cells morphologically resemble endothelial cells and show surface markers associated with both blood and lymphatic endothelium [Cancian et al., 2013; Carroll et al., 2004; Gessain and Duprez, 2005; Pyakurel et al., 2006]. The vast differences between KS and PEL begged the question of how a single virus could cause both diseases. The

initial idea was that the two malignancies resulted from KSHV infection of different cell types, with the B-cell-like PEL originating from transformed B-cells, and the endothelial-like KS originating from endothelial cells [Staskus et al., 1999]. The hypothesis that PEL originated from B-cells was consistent with the known ability of KSHV to establish latent infection of B-cells *in vivo*, and the fact that PEL cells closely resemble B-cells. None of the evidence suggested a different origin.

The origin of KS was less clear. First, KS spindle cells did not resemble any particular cell type. Morphologically they were similar to endothelial cells, but molecularly they displayed an indistinct phenotype; they expressed genes specific to blood endothelial cells, but also genes specific to lymphatic endothelial cells. In other words, it seemed clear that they originated from endothelial cells of some kind, but whether they came from lymphatic or blood endothelial cells remained unknown. This indistinct phenotype was later found to result from the virus reprogramming host cells. When the virus infected blood endothelial cells, it would push them to a more lymphatic-like gene expression profile [Carroll et al., 2004]. And when it infected lymphatic endothelial cells, it would push them to a more blood-like gene expression profile [Wang et al., 2004]. Regardless of which cell type they started out as, the reprogrammed cells showed a gene expression pattern that was roughly 70% lymphatic endothelial and 30% blood endothelial. The fact that KSHV reprogramed host cell gene expression patterns meant that those patterns could not be used to infer which type of cells gave rise to KS spindle cells.

After it was discovered that KSHV had the ability to reprogram surface marker protein patterns, the next logical steps were to determine if the virus was able to reprogram host cell morphology as well and, if so, what conditions were needed to trigger it, and by what mechanism was it accomplished. The vast majority of KS cases are the result of KSHV-infected individuals becoming immunocompromised. Immunosuppression triggers the viral genomes in latently infected cells to begin producing the viral protein RTA (Replication and Transcription Activator) [Greene et al., 2007; Guito and Lukac, 2012]. RTA is a transcriptional transactivator that stimulates the production of a cascade of viral proteins that initiate lytic replication. Somehow either RTA, or the proteins in the cascade it activates, cause KS to form. If

morphological reprogramming is part of KS formation, then RTA is either directly or indirectly (through the other genes it activates) responsible for that reprogramming.

Previous work in our lab tested whether lytic reactivation of KSHV was capable of reprogramming cell morphology *in vitro* (J. Chen, unpublished work, 2013). The PEL-derived BCBL-1 (Body Cavity Based Lymphoma 1) cell line was selected as a model because it was latently infected with KSHV. BCBL-1 cells were treated with phorbol ester (12-O-tetradecanoyl phorbol-13 acetate [TPA], Sigma, MO) to induce RTA expression and initiate lytic replication. At that time some of the cells were indeed morphologically reprogrammed to become spindle shaped and began adhering to each other and to the culture flask. Those reprogrammed cells (dubbed BCBL-E cells) were isolated and cultured. Whole transcriptome sequencing (RNA-Seq) comparing BCBL-E and BCBL-1 cells showed expression of over 500 genes had been either up- or down-regulated by at least two-fold; but gave no way to differentiate the effects caused directly by RTA from effects caused by the cascade of other viral proteins produced during lytic reactivation.

Differentiating those effects required the creation of a new set of KSHV-naïve cell lines modified to express RTA, but lacking all other viral components. To this aim, a plasmid vector coding for expression of RTA was transfected into a mammalian cell line that was not infected with KSHV. In this way RTA would be expressed in that cell line and would be capable of interacting with the host genome, but no other viral genes would be present. Therefore, if morphological reprogramming did occur in those cells, then it would have to have been caused directly by RTA. MDCK (Madin-Darby Canine Kidney) cells were chosen because they were not infected with KSHV, and unlike many human cell lines they were permissive of expressing KSHV proteins. Two types of MDCK cell lines were created by transfection with one of two plasmids. The treatment cells were named RTA+ and received a plasmid coding for constitutive expression of viral RTA. The control cells were named RTA- and received a similar plasmid that lacked the RTA gene. Like the BCBL-E cells, the RTA+ cells took on a spindle-like shape. RNA-Seq comparison between RTA+ and RTA- cells identified 180 genes for which expression was up- or down-regulated at least two-fold (J. Chen, unpublished work, 2014). With KSHV's ability to reprogram host cell morphology established, and immunosuppression-induced RTA production

identified as the trigger, the goal of this study was to identify the possible mechanisms RTA was using to effect that reprogramming.

References

Beral V, Peterman TA, Berkelman RL, Jaffe HW. 1990. Kaposi's sarcoma among persons with AIDS: a sexually transmitted infection? The Lancet 335:123–128.

Cancian L, Hansen A, Boshoff C. 2013. Cellular origin of Kaposi's sarcoma and Kaposi's sarcoma-associated herpesvirus-induced cell reprogramming. Trends in Cell Biology 23:421–432.

Carroll PA, Brazeau E, Lagunoff M. 2004. Kaposi's sarcoma-associated herpesvirus infection of blood endothelial cells induces lymphatic differentiation. Virology 328:7–18.

Cesarman E, Chang Y, Moore PS, Said JW, Knowles DM. 1995. Kaposi's Sarcoma–Associated Herpesvirus-Like DNA Sequences in AIDS-Related Body-Cavity–Based Lymphomas. New England Journal of Medicine 332:1186–1191.

Chang Y, Cesarman E, Pessin M, Lee F, Culpepper J, Knowles D, Moore P. 1994. Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. Science 266:1865–1869.

Curran JW, Jaffe HW. 2011. AIDS: the early years and CDC's response. MMWR Supplements 60:64–9.

Eltom MA, Jemal A, Mbulaiteye SM, Devesa SS, Biggar RJ. 2002. Trends in Kaposi's sarcoma and non-Hodgkin's lymphoma incidence in the United States from 1973 through 1998. Journal of the National Cancer Institute 94:1204–10.

Ensoli B, Stürzl M. 1998. Kaposi's sarcoma: a result of the interplay among inflammatory cytokines, angiogenic factors and viral agents. Cytokine & Growth Factor Reviews 9:63–83.

Farge D. 1993. Kaposi's sarcoma in organ transplant recipients. The Collaborative Transplantation Research Group of Ile de France. The European Journal of Medicine 2:339–43.

Gessain A, Duprez R. 2005. Spindle cells and their role in Kaposi's sarcoma. The International Journal of Biochemistry & Cell Biology 37:2457–2465.

Greene W, Kuhne K, Ye F, Chen J, Zhou F, Lei X, Gao S-J. 2007. Molecular biology of KSHV in relation to AIDS-associated oncogenesis. Cancer Treatment and Research 133:69–127.

Grundhoff A, Ganem D. 2004. Inefficient establishment of KSHV latency suggests an additional role for continued lytic replication in Kaposi sarcoma pathogenesis. Journal of Clinical Investigation 113:124–136.

Guito J, Lukac DM. 2012. KSHV Rta Promoter Specification and Viral Reactivation. Frontiers in Microbiology 3:1–21.

Hutt MSR, Burkitt DP. 1965. Geographical Distribution of Cancer in East Africa: A New Clinicopathological Approach. British Medical Journal 2(5464):719–722.

Patel S, Xiao P. 2013. Primary Effusion Lymphoma. Archives of Pathology & Laboratory Medicine 137:1152–1154.

Pyakurel P, Pak F, Mwakigonja AR, Kaaya E, Heiden T, Biberfeld P. 2006. Lymphatic and vascular origin of Kaposi's sarcoma spindle cells during tumor development. International Journal of Cancer 119:1262–1267.

Staskus KA, Sun R, Miller G, Racz P, Jaslowski A, Metroka C, Brett-Smith H, Haase AT. 1999. Cellular tropism and viral interleukin-6 expression distinguish human herpesvirus 8 involvement in Kaposi's sarcoma, primary effusion lymphoma, and multicentric Castleman's disease. Journal of Virology 73:4181–7.

Wang H-W, Trotter MWB, Lagos D, Bourboulia D, Henderson S, Mäkinen T, Elliman S, Flanagan AM, Alitalo K, Boshoff C. 2004. Kaposi sarcoma herpesvirus–induced cellular reprogramming contributes to the lymphatic endothelial gene expression in Kaposi sarcoma. Nature Genetics 36:687–693.

Chapter 2. KAPOSI'S SARCOMA-ASSOCIATED HERPESVIRUS REPLICATION AND TRANSCRIPTION ACTIVATOR REGULATES EXTRACELLULAR MATRIX SIGNAL PATHWAY

2.1. Abstract

Kaposi's sarcoma (KS) is a malignancy characterized by spindle-shaped tumor cells believed to be of endothelial origin. KS arises when individuals infected with Kaposi's sarcoma-associated herpesvirus [KSHV; also known as human herpesvirus 8 (HHV8)] become severely immunocompromised. Immunosuppression triggers production of KSHV's lytic switch protein RTA (Replication and Transcription Activator) in latently infected cells. RTA initiates lytic replication by setting off a cascade of both host and viral gene expression changes. Previous work in this lab showed that RTA was capable of reprogramming host cells to display a spindle-shaped morphology in both KSHV-infected and KSHV-naïve cell cultures *in vitro*. This study analyzed the gene expression changes in KSHV-naïve cells caused by RTA to identify possible mechanisms driving the morphological reprogramming. The analysis identified 8 genes involved in extra cellular matrix (ECM) remodeling, 1 gene involved in blocking cell-cell and cell-ECM adhesion, and 1 gene involved in epithelial/endothelial differentiation as the most likely drivers of the reprogramming.

2.2. Introduction

Kaposi's sarcoma-associated herpesvirus (KSHV; also called human herpesvirus 8 or HHV8) is a gammaherpesvirus which infects B-lymphocytes and endothelial cells [Dittmer and Damania, 2013]. KSHV can remain in its latent phase for decades in healthy individuals before immunosuppression triggers the production of the virus's lytic switch protein RTA (Replication and Transcription Activator) [Greene et al., 2007; Guito and Lukac, 2012]. RTA induced lytic reactivation of KSHV in severely immunosuppressed individuals directly causes nearly all cases of Kaposi's sarcoma (KS) and primary effusion lymphoma (PEL) [Moore and Chang, 2003]. Prior to widespread use of effective anti-retroviral therapy in the developed world, KS was found in approximately half of all AIDS patients [Moore and Chang, 2003], and it is currently the sixth most common cancer diagnosed in Africa [Abratt and Vorobiof, 2003].

KS tumors are highly vascular lesions made up of spindle-shaped cells, immature blood vessels, and inflammatory infiltrate [Ensoli and Stürzl, 1998]. Spindle cells display an endothelial-like morphology, and a pattern of surface marker proteins that is intermediate between blood and lymphatic endothelium [Pyakurel et al., 2006]. Reprogramming of surface marker patterns is not uncommon during oncogenesis, and KS is no exception. When either blood or lymphatic endothelial cells are infected with KSHV *in vitro*, the same intermediate surface marker pattern is induced [Carroll et al., 2004; Wang et al., 2004]. Additionally, spindle cells from biopsies often contain fragments of KSHV DNA as well as complete virions, but rapidly lose both when cultured, implying that the infection is not stable [Orenstein et al., 1997]. In contrast to KS, the B-lymphocyte-derived PEL cells do not form solid tumors, and are capable of maintaining a stable latent KSHV infection when cultured [Patel and Xiao, 2013]. Both morphologically and in regard to surface markers, PEL cells resemble B-cells growing rapidly in suspension. And while the progression of classic KS is so slow that patients often die of unrelated causes, patients diagnosed with PEL rarely survive a year. The most commonly accepted hypothesis for how a single virus can cause these two starkly different malignancies is that KS results from transformed endothelial cells, and PEL results from transformed B-cells [Staskus et al., 1999]. The many similarities between B-lymphocytes and PEL cells strongly support a B-cell origin. The origin of KS was less clear; given that the surface marker patterns in spindle cells were being reprogrammed by KSHV, the only evidence that spindle cells originated from endothelial cells was their endothelial-like morphology. However, previous work in our lab showed that RTA-induced lytic reactivation of KSHV was capable of morphologically reprogramming PEL-derived BCBL-1 (Body Cavity Based Lymphoma 1) cells *in vitro* to display a spindle-like morphology (J. Chen, unpublished work, 2013); indicating that (like surface marker patterns) morphology is not a reliable indicator of the cellular origin of KS spindle cells.

The finding that RTA-induced lytic reactivation of KSHV is capable of reprogramming host cell morphology opened the possibility that KS may share the same B-cell origin as PEL, but it came with two key limitations. First, that finding did not indicate whether the cause of the reprogramming was RTA itself or one of the many viral proteins produced in response to RTA. And second, it did not provide any insight into the mechanisms by which the reprogramming

was accomplished. The first point was addressed when further work in our lab showed that KSHV-naïve MDCK (Madin-Darby Canine Kidney) cells transfected with a plasmid vector coding for RTA expression adopted a spindle-shaped morphology, indicating that the reprogramming was being caused by RTA directly (J. Chen, unpublished work, 2014). This study aimed to address the second point by examining differences in gene expression between MDCK cells transfected with the RTA coding plasmid (RTA+ cells) and MDCK cells transfected with a similar plasmid lacking the RTA coding region (RTA- cells) in order to identify signaling pathways altered by RTA that could be responsible for the observed morphological reprogramming. A number of factors have been implicated in the control of cell morphology. The extracellular matrix (ECM) is one example. Changes in the ECM can result in dramatic changes to cell shape [Lukashev and Werb, 1998; Watt, 1986]. Individual ECM components such as collagen and laminin can modulate cell shape [Elliott et al., 2003; Ingber et al., 1995; Watt, 1986], as can large scale remodeling of the ECM [Gospodarowicz et al., 1978; Guilak et al., 2009; Lin and Bissell, 1993; Lukashev and Werb, 1998]. Other factors can also influence cell shape such as actin [Pollard and Cooper, 2009], neurotrophins [Bibel, 2000], and cell surface mechanics such as mechanical stress and cell-cell/cell-ECM adhesion [Ingber et al., 1995; Lecuit and Lenne, 2007]. We hypothesized that genes associated with one or more of these pathways would be altered by RTA.

2.3. Methods

2.3.1. Formal Virus Name

Family *Herpesviridae*, Subfamily *Gammaherpesvirinae*, Genus *Rhadinovirus*, Species *Human herpesvirus 8*.

2.3.2. Confirmation of Cell Line Identity

Cultures of the RTA+ and RTA- MDCK cell lines described in the introduction were utilized in this study. Presence or absence of the RTA gene in each cell line was verified in a three step process. First, DNeasy Blood and Tissue Kits (Qiagen, Hilden, Germany) were used to extract the

genomic DNA from BCBL-1 cells (BG, Figure 1). RNeasy Mini Kits (Qiagen, Hilden, Germany) were used to extract total RNAs from BCBL-1 cells and the RTA+ and RTA- cell lines. Total RNAs were reverse transcribed into cDNA from BCBL-1 cells (BC, Figure 1) and from the RTA+ and RTA- cells (RTA- and RTA+, respectively, Figure 1). Second, PCR was used to amplify the RTA gene and transcript from the cDNA extracted from the RTA+ and RTA- cells, as well as genomic DNA or cDNA from three control groups. The first two controls were gDNA and cDNA from a cell line (BCBL-1) infected with KSHV. The third control was the plasmid DNA used to create the RTA+ cells. The primers used in the PCR were capable of amplifying both the full length genomic loci of the RTA gene (~1350 bp) and the shorter cDNA version of the RTA transcript (~400 bp) that was used to construct the RTA+ plasmid. Third, an agarose gel electrophoresis was performed on the five PCR products to determine which versions of RTA (if any) were present in each sample. Presence of only the genomic-length version of RTA would indicate the presence of KSHV genome; presence of only the shorter cDNA-length version of RTA would indicate that the sample either contained RTA transcript (RTA+ in Figure 1), or contained the RTA+ plasmid; presence of both versions of RTA in sample would indicate that the cell line contained both KSHV genome and RTA transcript from active infection; and absence of both versions would indicate a lack of either RTA gene or RTA transcript.

2.3.3. Genes Eliminated for Technical Reasons

Prior to the beginning of this study whole-transcriptome sequencing (RNA-Seq) had generated a list of 180 genes whose expression level was at least 2-fold different in RTA+ cells as compared to RTA- cells. Before investigating potential biological significance, this study began by screening each gene to ensure that it met two technical criteria. The first criterion was that the gene must be expressed at some level in both the RTA+ and RTA- cells. Genes showing zero expression in either cell line were eliminated because fold-change cannot be computed if one of the values is zero, and the fact that RTA is a transcriptional transactivator (and therefore unlikely to turn a gene entirely on or off). The second criterion was that the gene must support cDNA specific primers. Primer sets were designed for each of the remaining genes using the

Primer-Blast tool from NCBI (http://www.ncbi.nlm.nih.gov/tools/primer-blast) with the condition that either the forward or reverse primer must span an exon-exon gap to ensure that genomic DNA would not be amplified (Table 1).

Table 1: Forward and reverse primer sequences used for qPCR. Note: Both the order of genes listed in this table and their placement in either the right or left column are arbitrary and do not signify any meaningful differences between either the genes or their primer sequences.

Gene	Primer Sequence
CCL2 - F	TCACCCAGCCAGATGCAATTA
CCL2 - R	TCTTTGGGACACTTGCTGCT
CTSS - F	AAATACCAGGGTTCTTGTGGTG
CTSS - R	AGCAATCTACCAAGTTCTGTGC
SNAI2 - F	GGACACGCACACAGTGATTATT
SNAI2 - R	CCATTGGGTAGTGGGGAGTG
GUCA2A - F	CCATGAACACCTTCCTGCTC
GUCA2A - R	AGAAGTCTCCGTCCTGCAC
PRKDC - F	CATTGCACCCGGAGATGAAC
PRKDC - R	TGCTCACACAGTCCTCCAAA
CXCL10 - F	CCACATGTTGAGATCATTGCCA
CXCL10 - R	TCAGACATCTTTTCTCCCCACT
FUT10 - F	GAGGTCTCCAATCAGCGACT
FUT10 - R	GTAAGCCCTTTTCCTGAAGCC
SEPP1 - F	AGGTTTCAGAGCACATTCCTG
SEPP1 - R	AGACGGCCACATCTGTCATA
ADRA1B - F	GGCATTGTGGTCGGTATGTT
ADRA1B - R	TTGAAGTAGCCCAGCCAGAA
CTSC - F	CCAGTATGCTCAAGGCTGTG
CTSC - R	ACGGAAGCAGTCATTTGGTT
BCL2L2 - F	AACCTGGCAGGGCTCCAC
BCL2L2 - R	ACTAGAGCCCGTGTGTCTGG

Gene	Primer Sequence
SETMAR - F	CCCTTCCAGTATACTCCTGACC
SETMAR - R	ATGACGGAGACAGGAGCAAG
MGMT - F	TGTCCAGAGGAAATGACGGA
MGMT – R	TGAACGAACCTCGCTGGAAA
SLC1A1 - F	GACGGGACTGCACTCTATGA
SLC1A1 - R	CCGTGACACTGATGGTGATG
ZFPM2 - F	CCGTGACACTGATGGTGATG
ZFPM2 - R	GCTCAGATTTTCAGGCCCAA
CXCR7 - F	TTGCTCACGACGGGATCATT
CXCR7 - R	GAACGGCCACATTGAGACTTT
CYP4A37 - F	CAAACCCAGAGGTGTTCGAC
CYP4A37 - R	CTCGTTCATGGCAAAGTGCT
COL1A2 - F	TACAGGAGGCAACTGCAAGA
COL1A2 - R	CTGGGATACCATCGTCACCA
PDPN - F	GAGAAAGATGGCCTGGCAAC
PDPN - R	GCTCTTTAGGGCGAGTACCT
PSMB8 - F	GGGAGCGTATCTCAGTGTCA
PSMB8 - R	CCAGGACCCTTCTTATCCCA
MMP9 - F	GTGTTAGGGAGCACGGAGAT
MMP9 - R	CCCTTGCCCAGAGTCCATAA
CYP4A38 - F	AGCACAGTTCTACTGGCACA
CYP4A38 - R	CCCGACCCTTTTGGAACTCT

Table 1 continued

Gene	Primer Sequence
CYP4A11 - F	GCTGCAGCTGCTACTGAAAT
CYP4A11 - R	CTGGTTTTGGGTCAGATCGC
ARG2 - F	AAGGTCAGTGGGCTAAAACG
ARG2 - R	AAAGGAAGGAGTCTGTTCAGGA
PLAU - F	TGTCTAGCGCATCAAACTGC
PLAU - R	GAGTGACCATTCCCCTCGTA
KRT19 - F	GAACCACGAGGAGGAAATCAG
KRT19 - R	TCACTCAGGATTTTGGCAAGG
PLAU - F	GTCTACCTGGGTCGGTCAAA
PLAU - R	CGGATCTTCAGCAAGGCAAT
MT1E - F	CAAGAAGAGTTGCTGCTCCTG
MT1E - R	GGAACAGGTGTTCTCACACATC
CD47 - F	GGTGTTTGGATGTCTCCTTTCA
CD47 - R	CACTGGGGTGCACTCTGA
SELT - F	CAAGCTCCTAGCATCTGGCA
SELT - R	ACCACACAGGCACATCATTC
RPS10 - F	CAGGCGCCAAGATGTTGAT
RPS10 - R	CACATTCTTGTCGGCCAGTT
LAMC2 - F	TGGATTCCGCTGTCTCAACT
LAMC2 - R	GGCGCTAAGAGAACCTTTGG
RPS15A - F	GAATGTCCTGGCAGATGCTC
RPS15A - R	AAATTCGCCAATGTAACCATGC
MAP2K1 - F	GTGGGGAGATCAAGCTCTGT
MAP2K1 - R	TGGAGTCTTTCTGGCGACAT
DSG3 - F	GGCCATTGATGAAAACACAGG
DSG3 - R	GCTGAGACAACCACAGAACG
PODXL - F	CGGGGAACAATTCGGATGAC
PODXL - R	AGAGGTTCGTCTGGACAGTG

Gene	Primer Sequence
ARG2 - F	ACCCCTCACCACTTCATCTG
ARG2 - R	TGGGAGTTGTGGTACCCATT
LMAN2 - F	ACTGGTGACCTGTCTGACAAT
LMAN2 - R	ACACTGGGCTCAATCTTGGT
EDN1 - F	GTCAACACTCCTGAGCACATT
EDN1 - R	TGTGGTCTGTTGCCTTTGTG
CSF2 - F	GACGTGACTGCTGTGATGAATA
CSF2 - R	AAGGGATTCTTGAGGCTGGT
RPL27A - F	CAACATGCCATCCAGACTGAG
RPL27A - R	CATTACCCCGGCCTCCTG
RAB2A - F	CCATGGCGTACGCTTATCTC
RAB2A - R	AGCACCAAACTCTACACCGAT
DAG1 - F	GAACTCGTCAAGGTGTCAGC
DAG1 - R	ATCGGTGTCAAGAGGGAGAC
TIMP1 - F	CAAATCGTCATCAGGGCCAA
TIMP1 - R	TGTCGACGAAGCGGATGT
HSPB8 - F	AAATCCAGCTTCCCGCAGA
HSPB8 - R	GGGGAAGCTCATTGTTGAAGT
TXNRD1 - F	ACCGAGGAGACGGTTAAACA
TXNRD1 - R	AGGACCCACAAACTGTCCAT
CPNE1 - F	TCCATTTCCTGTGACCACCT
CPNE1 - R	CAGTTCGGCCAAGCTCAG
PTGER2 - F	TGCTCCTTGCCTTTCACAATTT
PTGER2 - R	AGGAGGCCTAAAGATGGCAA
RPS24 - F	AGACATGGCCTGTATGAGAAGA
RPS24 - R	CCAATCTCCAGCTCACTTCTTT
AP3S2 - F	ATCTTGGACCTCATCCAGGTTT
AP3S2 - R	TTCCAATACCATCCCACCCA

Table 1 continued

Gene	Primer Sequence
GUSB - F	ACTACTTCAAGACGCTGATTGC
GUSB - R	ACAGATGACGTCCACATACGG
PLA2G7 - F	GGCGCCCGGAAGTTTAAG
PLA2G7 - R	GGATGAACCAGTGTGAGGCA
GSTA4 - F	GACTCCAGAAAGCCTGAAAGG
GSTA4 - R	ACTCCAGCGGCAGCTAAAA
CD97 - F	CCTGGAGAATCCTCAGCCAA
CD97 - R	GCAGGTGGTGCTGTTATTCC
POLDIP2 - F	CTCTCGTCCCGAAACCGA
POLDIP2 - R	CAGGACGACACCTCGGTAG
LGALS9 - F	ACAGCTGCGATTTCAAGGTG
LGALS9 - R	GCCTGGAGACTGGAAGCTAA
RBM47 - F	CTCTGAGGGTGGGACACAAG
RBM47 - R	CGGAGACCTGCCAAACAC
FASTK - F	CGCTGCAAGTACAGTCACAA
FASTK - R	AGGAAGTCTGTGCAGTAGCC
SLC35C1 - F	CTCACTCACCACCGTCTTCA
SLC35C1 - R	AGCCAGAAGCCACCAATGAT
SEPX1 - F	TTTGAGCCGGGTGTCTATGT
SEPX1 - R	TCTGCATGAATGGTCTCCGT
PLA2G15 - F	TGGACCCCAGCAAAAGTAGT
PLA2G15 - R	TAGGGCCCGTTTTCATTTGG
PRKACA - F	AGGGCTACATTCAGGTGACA
PRKACA - R	GTCCACAGCCTTGTTGTAGC
APP - F	CGGTCCCAGGTTATGACACA
APP - R	TGTCATCCGAGTAGTTTTGCTC
C3H4orf52 - F	GCCGGAGGAGTACCAGAAAG
C3H4orf52 - R	TCGAGGTGGTGATAGCCTTC
SLC35A3 - F	TTGTACAGTGGCCCTCAGATT
SLC35A3 - R	GTAAACCCCAGCAAAACCACT
SNRPB2 - F	AAAATGCGTGGCACTTTTGC
SNRPB2 - R	TGCATTTGGTGTTCCCTGAC
TMEM47 - F	CAGCAGCGATTGGCAGATT
TMEM47 - R	TGTAGAAGCGCCTTCGAGAT
COX16 - F	GTGCGCGAAAAGCTGACAT
COX16 - R	TCCAACAATCAGCAACAGCA
NACA - F	CAGTCAATTCAGGCCTCCTTTT
NACA - R	AGATTCTGTTCCAGACCCTGT
TBX2 - F	ATGGGCATGGGTCACCTG
TBX2 - R	TTGGGATTCCCTGAGATGCC
GNA11 - F	ATATGACCAAGTGCTGGTGGA
GNA11 - R	AGAGGATAACCGAGGAGTTCTG
TIMP2 - F	TGCACATCACCCTTTGTGAC
TIMP2 - R	CGCGTGATCTTGCACTCAC
ANXA13 - F	TCGTCATAGCCAGTCTTACTCC
ANXA13 - R	ATGGCTTTTCGCTTTGGCA
H19 - F	ATAGGACATGGTCCGGTGTG
H19 - R	TTCAGGAACGCAAAGGAACG
RHPN2 - F	GATGGACCTGAGACAGGCTT
RHPN2 - R	GAAGCGGCTCTCCACAAATC
KRT14 - F	GTCCTCCAGAGATGTGACCT
KRT14 - R	GAGCAGCCTCAGTTCTTGGT
KDM5C - F	GGTGGCTGGTCTACAGATGA
KDM5C - R	AGAATACCCTTGACATCCCCAA
FLCN - F	CCCGGAAAACTTCAAGCCTTA
FLCN - R	AGCAGAGGGCCGTTATATTCA

Table 1 continued

Gene	Primer Sequence
EFNB2 - F	CTGGGTCAGCCAGGCATAAT
EFNB2 - R	GTGCTGGAACCTGGATTTGG
NFKBIL1 - F	CTGGGATTCTGCTGAAGAGGA
NFKBIL1 - R	GGAAGCGTCATCTTCAAACCTC
RPS5-F	TCACTGGTGAGAACCCCCT
RPS5-R	CCTGATTCACACGGCGTAG
RPS19-F	CCTTCCTCAAAAAGTCTGGG
RPS19-R	GTTCTCATCGTAGGGAGCAAG
RPL8-F	CCATGAATCCTGTGGAGC
RPL8-R	GTAGAGGGTTTGCCGATG
DLA-64-F	CCCTGCAGCTCACAGATCCT
DLA-64-R	CTGCCAGGTCAGGGTGATCT
DLA-88-F	TCTCATGCTGCACGTGATGA
DLA-88-R	ATCTTGCATCGCTCAGTCCC
DLA-12-F	GGCCGGGTCTCACACCTT
DLA-12-R	AGCGCAGGTCCTCGTTCA
HLA-DRB1 - F	AGGGACACCCCACCACATT
HLA-DRB1 - R	AACTCCTCCCGGTTATGGATG
EGF – F	CCAGAGCCAGGGTCAGTAAA
EGF – R	AGAGCCGGTGCCTCTATTAC
FGFR1OP2 – F	GCAGCACTCCAAGGAATTACA
FGFR1OP2 – R	CTTGCAACCCTGTTGCTCAT
FOSB – F	TCCAGGCGGAGACAGATCA
FOSB – R	AACTCCAGACGCTCCTTCTC
POFUT1 – F	GCACAGACCACTCCAGAAGTA
POFUT1 – R	AGGCGTTCTTCCAGTCAGA

Gene	Primer Sequence
RHBDF2 – F	GAAGGACTGCTCGGAGACTT
RHBDF2 – R	CTGGCCCAGATCAGACTTGT
TES - F	GCTGGCTGCGATGAGTTAATA
TES - R	AGCTAGGATGTTGTCGCAGT
FZD6 - F	CCCAGGTCAAGAGAACAGGA
FZD6 - R	TGGAGTTACCCTTCCTTCACT
BRCA1 - F	AGTGGCTTCCATGCGATTG
BRCA1 - R	CTGCAGCAGTTCTGGGAATC
MUC1 - F	ACTGTTCCACCTCCTCCCA
MUC1 - R	CTCCCTGTGCTGTAAGCTCT
CLNS1A - F	TCTGTTGCTGAAGAAGAAGACAG
CLNS1A - R	GAACATTGCCTCCAATGCTGA
CGN - F	GAGGAGCTTGGGGAGAAGAT
CGN - R	TCTGGCACAGCTCCTTCTTA
ERBB2 - F	TCATTGCTCACAGCCAAGTG
ERBB2 - R	TTCAGGATCTCTGTGAGGCTT
EBAG9 - F	TATGCAGATCTGGCAGAGGAC
EBAG9 - R	CCCAGGAAGTCCACTCTTCA
TGFB1 - F	ATGTCACTGGAGTCGTGAGG
TGFB1 - R	GGCTGGAACTGAACCCGTTA

2.3.4. Analysis for Potential Biological Significance

The next step was to determine which of the gene expression changes identified by RNA-Seq were potentially biologically significant. This was done in three steps. First, the complete list of all genes meeting the previously stated technical requirements was analyzed using both GeneAnalytics™ (http://geneanalytics.genecards.org) and Ingenuity Pathway Analysis® (http://www.qiagen.com/ingenuity) software in order to identify associated pathways. Next, those pathways were screened for biological significance and/or potential association with the observed morphological changes using GeneCards™ (http://www.genecards.org/) and the NCBI gene information database (http://www.ncbi.nlm.nih.gov/gene). The results of those software analyses are presented in appendices A and B. Finally, a literature search was performed for each of the remaining genes and pathways in order to identify any additional associations.

2.3.5. qPCR Verification of RNA-Seq Results

The relative expression levels reported by RNA-Seq of genes flagged as potentially biologically significant was verified by quantitative real-time PCR (qPCR). To this aim, RNeasy Kits (Qiagen, Hilden, Germany) were used to extract total RNA from RTA+ and RTA- cells. The isolated RNA was treated with DNase to degrade any genomic DNA present, and then reverse transcribed to create cDNA. That cDNA was then used to run a series of qPCR's using a StepOnePlus Real-Time PCR System (Life Technologies, Carlsbad, CA). Two to seven runs were conducted for each gene, and each run contained two replicates per gene. The difference between the mean Ct value (the number of cycles required for a well to cross the threshold of detection) for RTA+ and RTA- for each gene was defined as ΔCt value for that gene. The ΔCt value for each gene was then adjusted by subtracting the ΔCt value for the housekeeping genes RPS5 and RPS9 to generate the ΔΔCt value for that gene. The fold change in expression for each gene was calculated as $2^{(-\Delta\Delta Ct)}$ for each run; with final results for each gene reported as the mean and standard deviation of the fold change values for all runs. The results were analyzed using the companion software for the StepOnePlus real time PCR system.

2.4. Results

2.4.1. Confirmation of Cell Line Identity

DNA was extracted from the RTA+ and RTA- cell lines along with DNA from plasmid used to create the RTA+ cells and genomic DNA (gDNA) and cDNA samples from KSHV infected BCBL-1 cells as positive controls. The RTA gene was amplified by PCR with primers capable of targeting both the genomic loci of the gene present in the virus itself and the shorter cDNA version of the gene used to create the RTA+ plasmid. An agarose gel run on the five PCR products (Fig 1) showed that, as expected, RTA- cells lacked either version of the RTA gene; and that RTA+ cells contained only the shorter cDNA version.

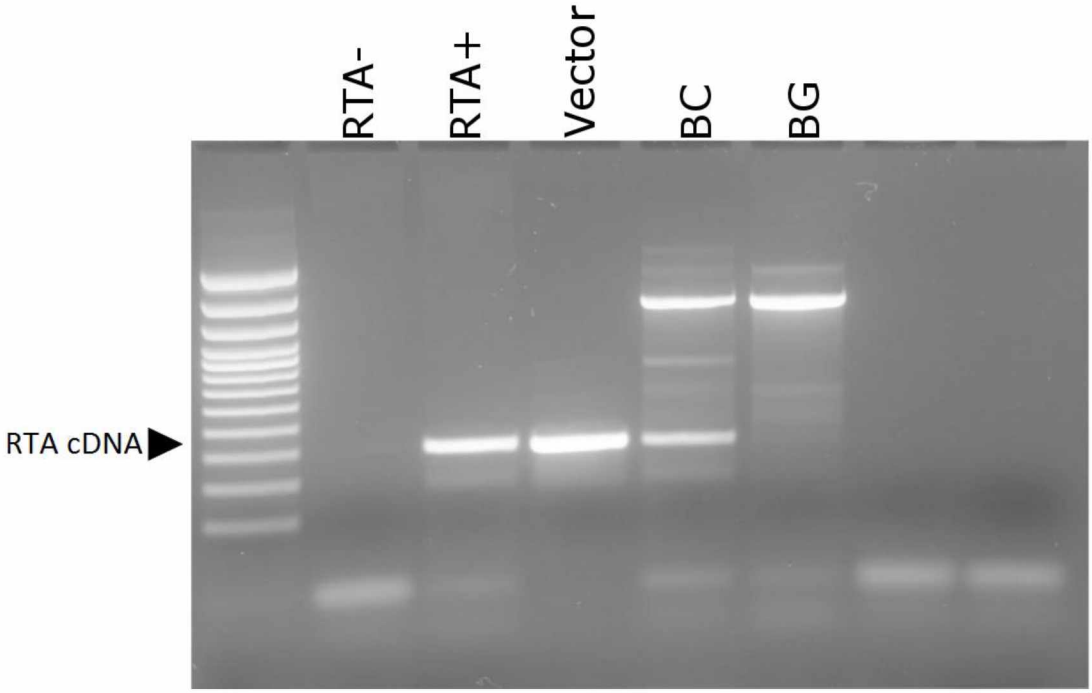

Fig 1: Plasmid uptake and RTA cDNA presence were verified by PCR and agarose gel electrophoresis. The plasmid used to make the RTA+ cell line (Vector), as well as both gDNA (BG) and cDNA (BC) from KSHV infected BCBL-1 cell cultures were used as positive controls to show the longer gDNA version of the RTA gene as well as the shorter cDNA version. RTA+ does not show the gDNA version, which indicates that it is not infected with KSHV; but it does show the cDNA version, indicating that it took up the plasmid. RTA- lacks both versions, indicating that it neither contains RTA cDNA, nor is it infected with KSHV.

2.4.2. Gene Selection

Previously completed RNA-Seq had identified a list of 180 genes of which expression was at least 2-fold different in RTA+ compared to RTA- cells (Figs 2-5). Of these 180 genes, 74 were eliminated because they were present in only one of the cell lines. The design of cDNA-specific primers for the remaining 106 genes resulted in the elimination of an additional 9 genes, for which primers spanning an exon-exon gap could not be constructed, leaving a total of 97 genes to be analyzed for biological significance.

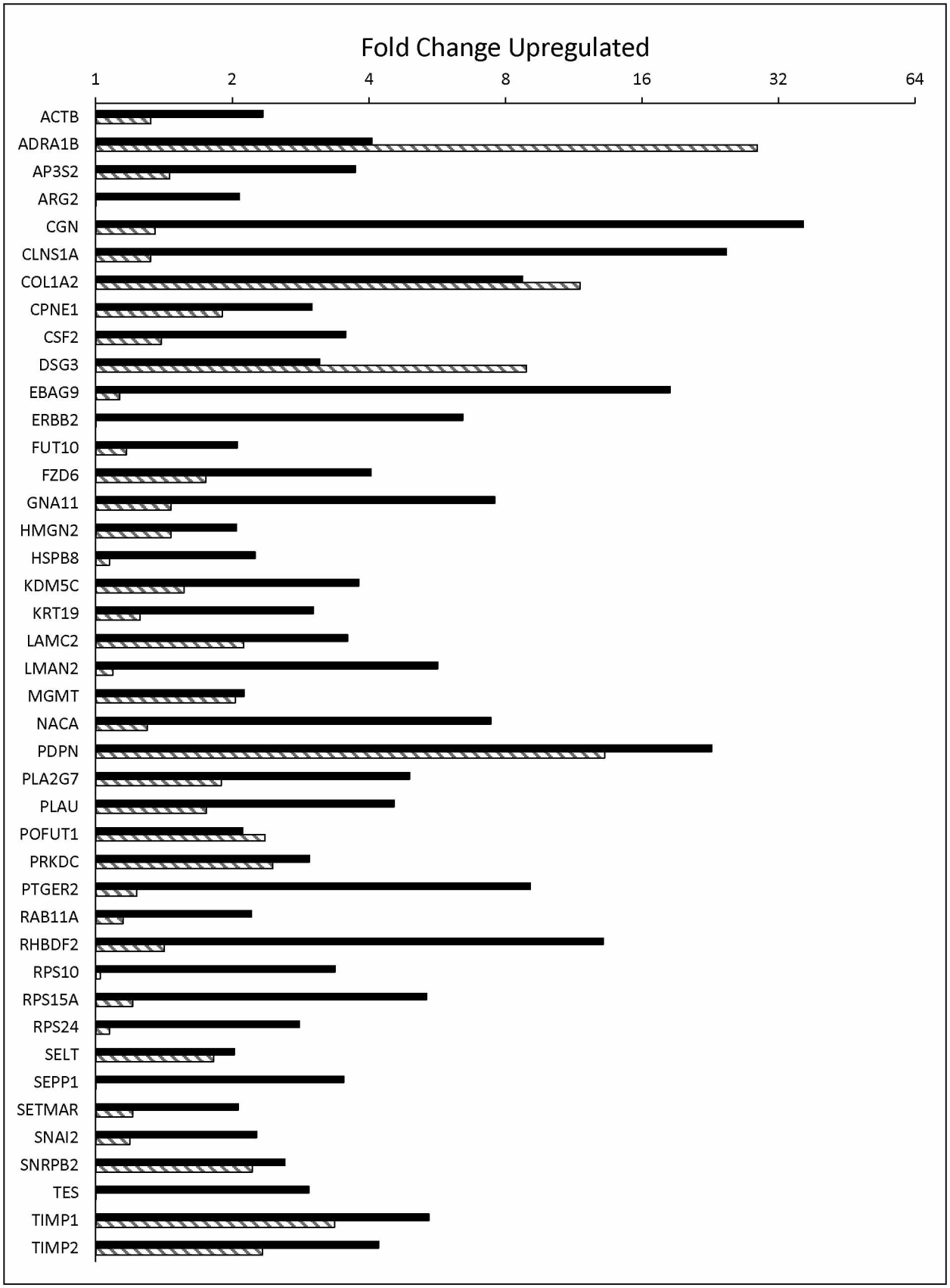
Fold Change Upregulated
1
2
4
8
16
32
64
ACTB
ADRA1B
AP3S2
ARG2
CGN
CLNS1A
COL1A2
CPNE1
CSF2
DSG3
EBAG9
ERBB2
FUT10
FZD6
GNA11
HMGN2
HSPB8
KDM5C
KRT19
LAMC2
LMAN2
MGMT
NACA
PDPN
PLA2G7
PLAU
POFUT1
PRKDC
PTGER2
RAB11A
RHBDF2
RPS10
RPS15A
RPS24
SELT
SEPP1
SETMAR
SNAI2
SNRPB2
TES
TIMP1
TIMP2

Fig 2: Comparison of selected RNA-Seq (black bars) and qPCR (striped bars) gene expression fold change measurements in RTA+ vs RTA- cell lines. This figure shows all genes that were both identified by RNA-Seq as upregulated by at least two-fold, and whose expression was subsequently measured by qPCR. Genes lacking a striped bar (ARG2, ERBB2, SEPP1, and TES) were shown to be upregulated by RNA-Seq, but subsequent qPCR measurements showed them to be downregulated. Note: due to the nature of RNA-Seq, error bars are not used when presenting RNA-Seq data.

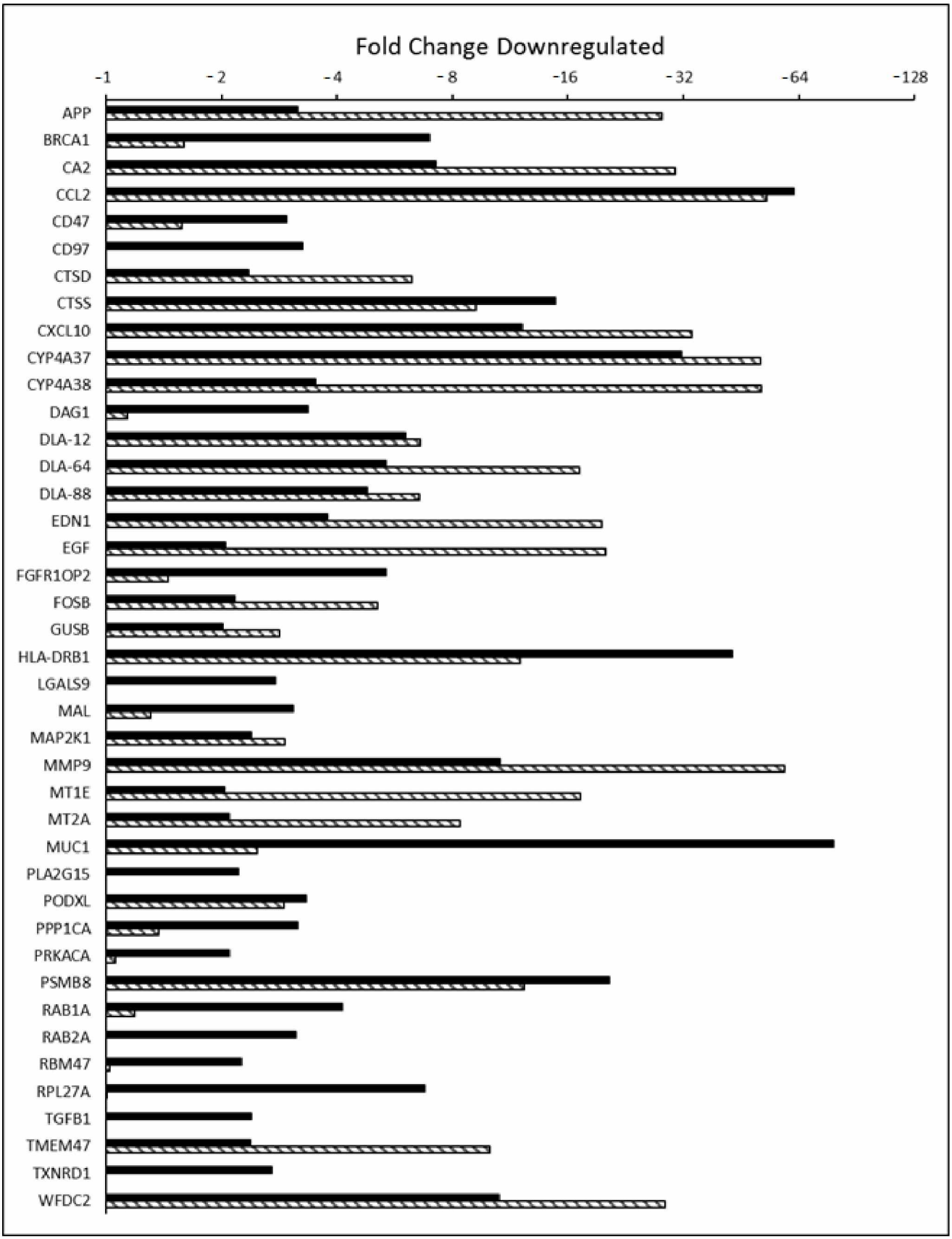
Fold Change Downregulated
-1
-2
-4
-8
-16
-32
-64
-128
APP
BRCA1
CA2
CCL2
CD47
CD97
CTSD
CTSS
CXCL10
CYP4A37
CYP4A38
DAG1
DLA-12
DLA-64
DLA-88
EDN1
EGF
FGFR1OP2
FOSB
GUSB
HLA-DRB1
LGALS9
MAL
MAP2K1
MMP9
MT1E
MT2A
MUC1
PLA2G15
PODXL
PPP1CA
PRKACA
PSMB8
RAB1A
RAB2A
RBM47
RPL27A
TGFB1
TMEM47
TXNRD1
WFDC2

Fig 3: Comparison of selected RNA-Seq (black bars) and qPCR (striped bars) gene expression fold change measurements in RTA+ vs RTA- cell lines. This figure shows all genes that were both identified by RNA-Seq as downregulated by at least two-fold, and whose expression was subsequently measured by qPCR. Genes lacking a striped bar (CD97, LGALS9, PLA2G15, RAB2A, TGFB1, and TXNRD1) were shown to be downregulated by RNA-Seq, but subsequent qPCR measurements showed them to be upregulated. Note: due to the nature of RNA-Seq, error bars are not used when presenting RNA-Seq data.

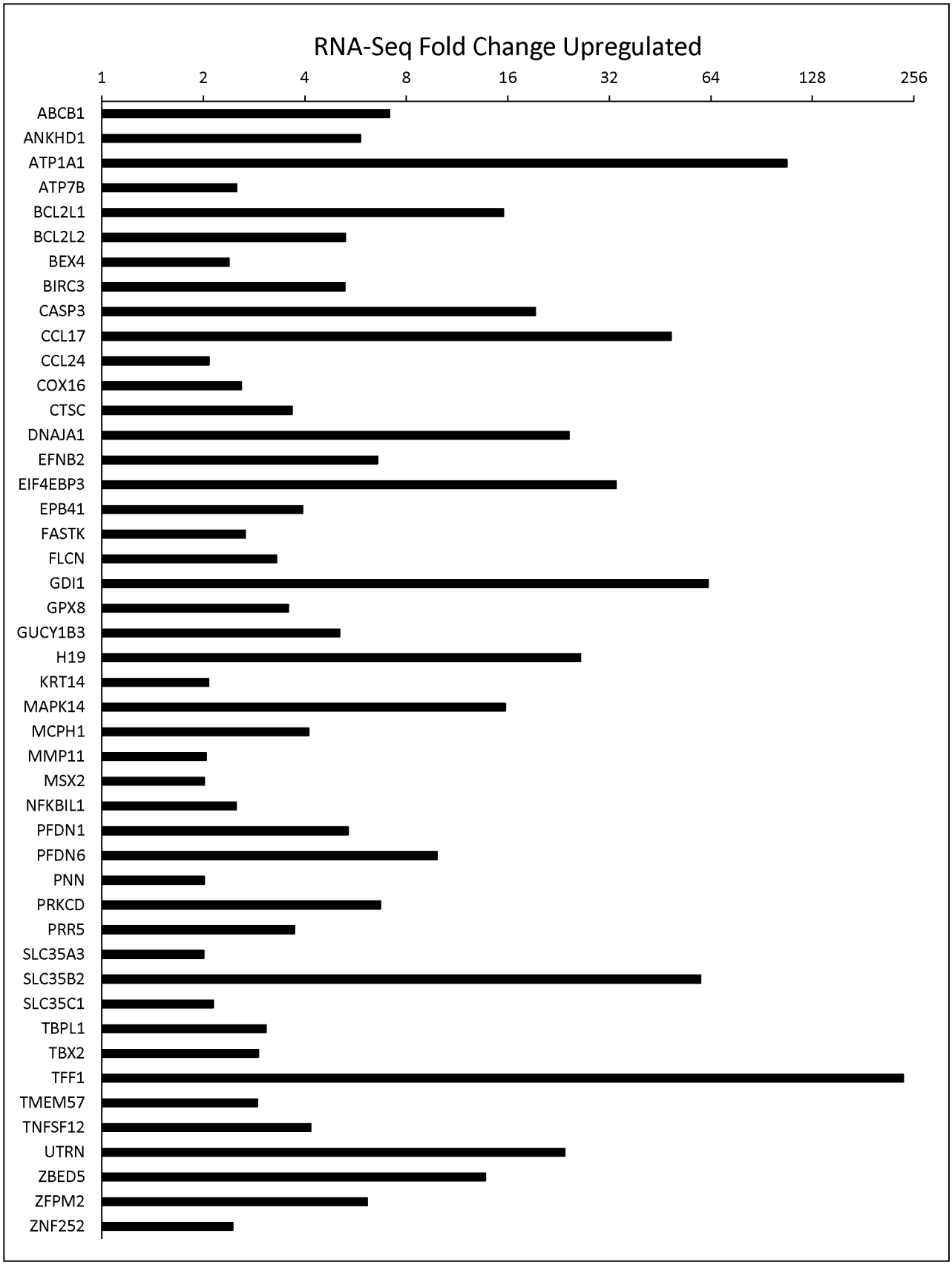
RNA-Seq Fold Change Upregulated
1
2
4
8
16
32
64
128
256
ABCB1
ANKHD1
ATP1A1
ATP7B
BCL2L1
BCL2L2
BEX4
BIRC3
CASP3
CCL17
CCL24
COX16
CTSC
DNAJA1
EFNB2
EIF4EBP3
EPB41
FASTK
FLCN
GDI1
GPX8
GUCY1B3
H19
KRT14
MAPK14
MCPH1
MMP11
MSX2
NFKBIL1
PFDN1
PFDN6
PNN
PRKCD
PRR5
SLC35A3
SLC35B2
SLC35C1
TBPL1
TBX2
TFF1
TMEM57
TNFSF12
UTRN
ZBED5
ZFPM2
ZNF252

Fig 4: Selected RNA-Seq gene expression fold change measurements in RTA+ vs RTA- cell lines. This figure shows all genes that were identified by RNA-Seq as upregulated by at least two-fold, but for which expression was not subsequently measured by qPCR either because suitable primer sets could not be designed, or because software analysis identified the genes as unlikely to be biologically significant. Note: due to the nature of RNA-Seq, error bars are not used when presenting RNA-Seq data.

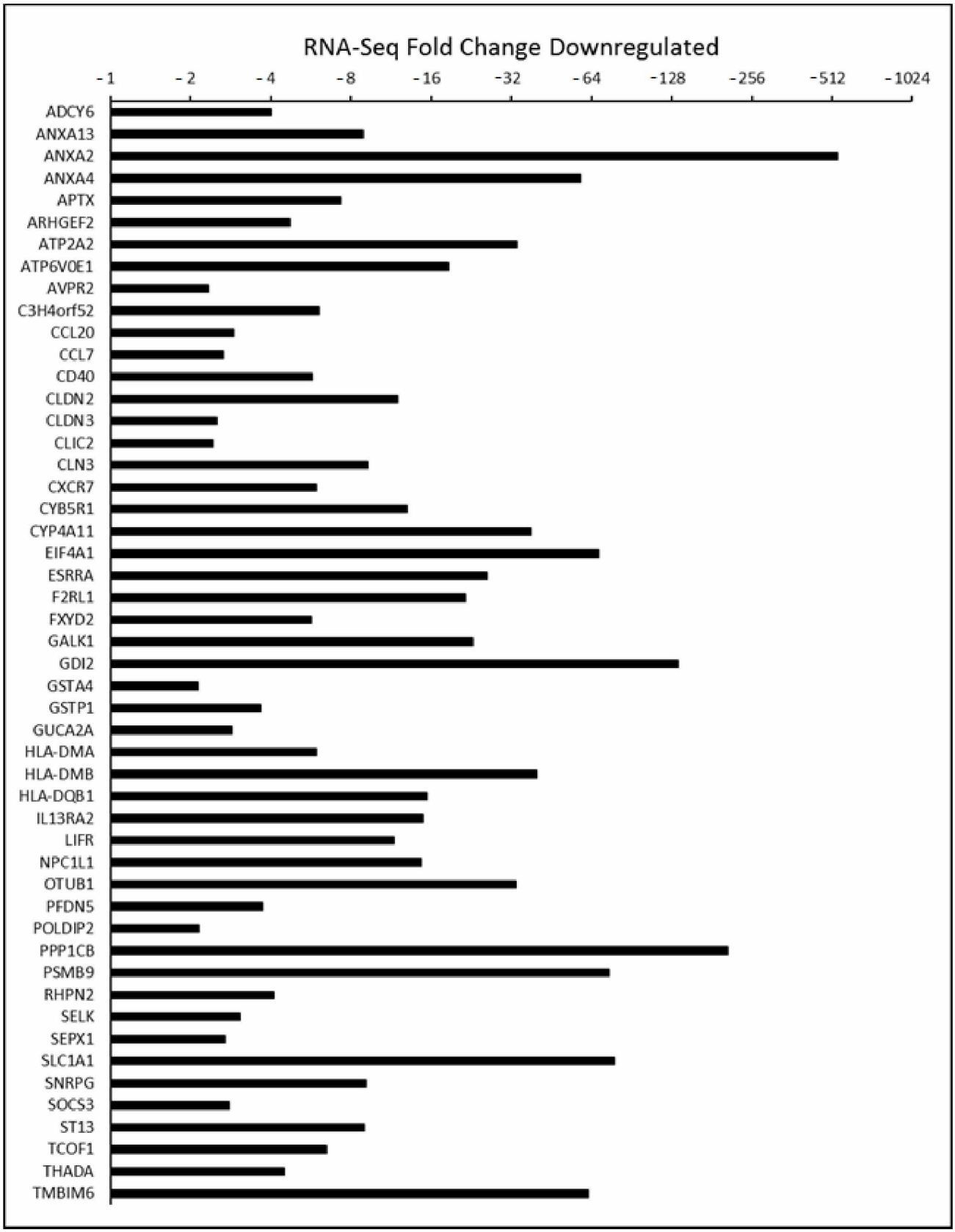
RNA-Seq Fold Change Downregulated
-1
-2
-4
-8
-16
-32
-64
-128
-256
-512
-1024
ADCY6
ANXA13
ANXA2
ANXA4
APTX
ARHGEF2
ATP2A2
ATP6V0E1
AVPR2
C3H4orf52
CCL20
CCL7
CD40
CLDN2
CLDN3
CLIC2
CLN3
CXCR7
CYB5R1
CYP4A11
EIF4A1
ESRRA
F2RL1
FXYD2
GALK1
GDI2
GSTA4
GSTP1
GUCA2A
HLA-DMA
HLA-DMB
HLA-DQB1
IL13RA2
LIFR
NPC1L1
OTUB1
PFDN5
POLDIP2
PPP1CB
PSMB9
RHPN2
SELK
SEPX1
SLC1A1
SNRPG
SOCS3
ST13
TCOF1
THADA
TMBIM6

Fig 5: Selected RNA-Seq gene expression fold change measurements in RTA+ vs RTA- cell lines. This figure shows all genes that were identified by RNA-Seq as downregulated by at least two-fold, but for which expression was not subsequently measured by qPCR either because suitable primer sets could not be designed, or because software analysis identified the genes as unlikely to be biologically significant. Note: due to the nature of RNA-Seq, error bars are not used when presenting RNA-Seq data.

2.4.3. Analysis for Potential Biological Significance

Software analysis using GeneAnalytics™ (http://geneanalytics.genecards.org) and Ingenuity Pathway Analysis® (http://www.qiagen.com/ingenuity) software identified all pathways associated with one or more of the 97 remaining genes (Appendices A and B). Additionally, IPA Biomarker Analysis identified 25 of the 97 genes as being associated with b-cells, epithelial cells, or both (Fig 6).

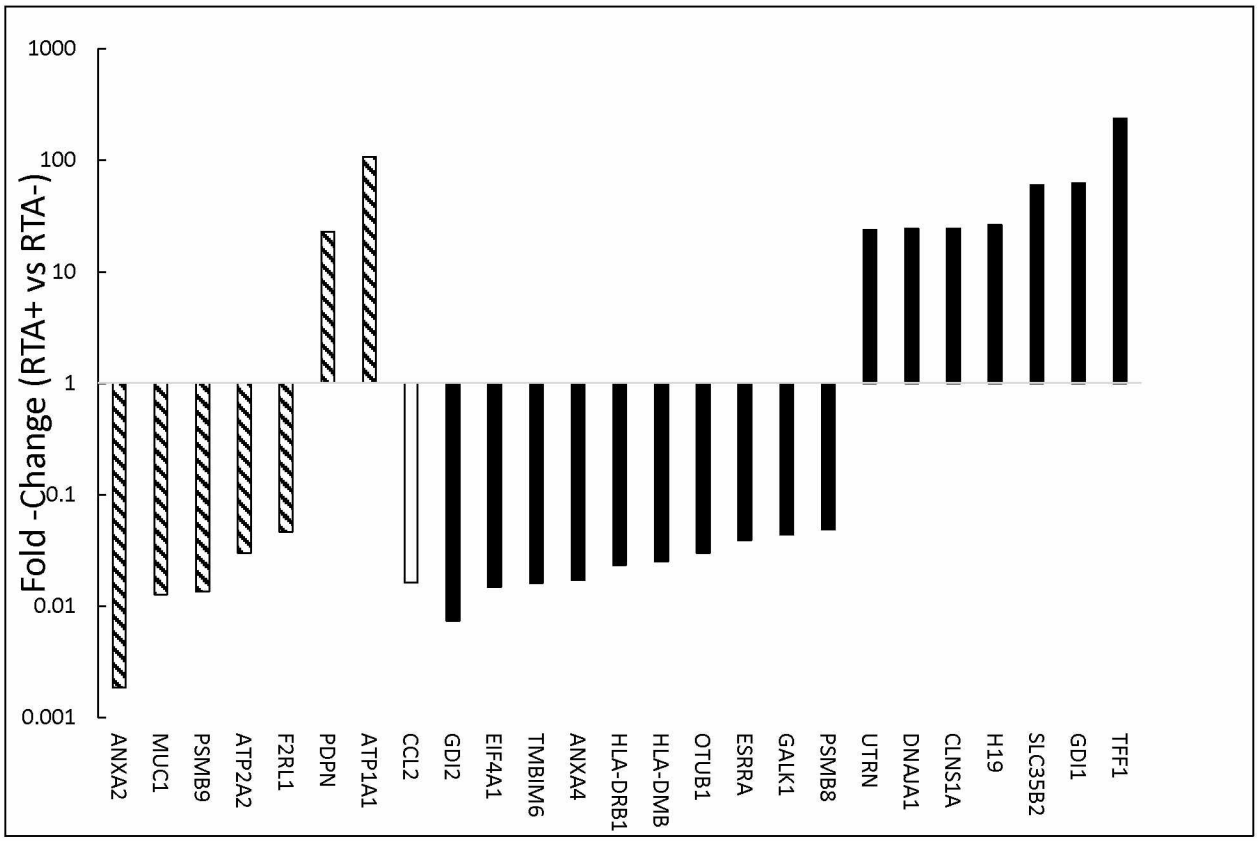

Fig 6: Results of IPA Biomarker Analysis on expression level differences between RTA+ and RTA- cell lines as determined by RNA-Seq. Striped bars represent genes associated with both epithelial cells and b-cells; the white bar represents a gene associated with epithelial cells, but not associated with b-cells; and the black bars represent genes associated with b-cells, but not with epithelial cells. Note: due to the nature of RNA-Seq, error bars are not used when presenting RNA-Seq data.

The pathways identified by the software analysis were then screened for biological significance and/or potential association with the observed morphological changes using GeneCards™ (http://www.genecards.org/), the NCBI gene information database (http://www.ncbi.nlm.nih.gov/gene), and reviewing available literature on each gene and pathway. Genes determined to be potentially biologically significant are listed in Table 2.

Table 2: Summarized results of literature review. Genes identified as upregulated by both RNA-Seq and qPCR are marked with green 'Up'. Genes identified as downregulated by both RNA-Seq and qPCR are marked with red 'Down'. Genes with a single asterisk (ANXA2, BCL2, CXCR7, PRKCD, & UTRN) indicate that qPCR was not performed. Genes with two asterisks (ERBB2 & TGFB1) indicate that RNA-Seq identified them as being up- or downregulated, but qPCR showed no difference in expression between RTA+ and RTA- cells. Neither BCL2 nor TNFα were themselves identified by RNA-Seq as being differentially expressed in RTA+ cells compared to RTA- cells, but were nonetheless included in the literature review. BCL2 was included because two similar genes [BCL2-like protein 1 (BCL2L1); and BCL2-like protein 2 (BCL2L2)] were identified by RNA-Seq as upregulated; and much more information is available about BCL2 than either BCL2L1 or BCL2L2. Similarly, TNFα was included because RNA-Seq identified the closely related TNFSF12 (TNF superfamily member 12) as upregulated in RTA+ cells.

Gene	Change in RTA+ Cells	Effect	References
ANXA2	Down*	Deregulation associated with metastasis; cytoskeletal attachment and organization; endocytosis and exocytosis	[Gerke and Moss, 2002; Hayes et al., 2004; Leal et al., 2015]
BCL2	(Up)*	Overexpression blocks apoptosis	[Anderson et al., 2014]
BRCA1	Down	Tumor suppressor gene involved in DNA repair	[Mersch et al., 2015]
CCL2	Down	Also called monocyte chemoattractant protein-1 (MCP-1); chemoattractant for natural killer cells, monocytes, and T lymphocytes; proinflammatory; proangiogenic; KSHV induced over-expression of CCL2 is accompanied by capillary-like structure formation	[Caselli et al., 2006; Conti and Rollins, 2004; Ohta et al., 2002]
CD47	Down	Regulates cell-cell adhesion in MDCK cells by reorganizing the actin cytoskeleton; important signaling molecule in the tumor microenvironment	[Shinohara et al., 2006; Sick et al., 2012]

Table 2 continued

Gene	Change in RTA+ Cells	Effect	References
CGN	Up	Involved in regulating gene expression during epithelial differentiation; regulates RHO GTPases; regulates claudin expression in MDCK cells; directly interacts with the actin cytoskeleton	[Aijaz et al., 2005; Citi et al., 1989, 2009, 2012; Guillemot and Citi, 2006; Guillemot et al., 2004; Matter and Balda, 2003]
COL1A2	Up	Component of the extracellular matrix protein type 1 collagen	[Elliott et al., 2003]
CTSD	Down	Overexpressed in breast cancer cells; degrades the extracellular matrix, thereby releasing ECM-bound basic fibroblastic growth factor	[Benes et al., 2008; Briozzo et al., 1991; Cavailles et al., 1993; Laurent-Matha et al., 2005; Liaudet-Coopman et al., 2006; Rochefort et al., 1989; Takei et al., 1997]
CTSS	Down	Knockout impairs tumor formation and angiogenesis; necessary for MHC class II peptide processing	[Gocheva et al., 2006; Shi et al., 1999; Wang et al., 2005]
CXCL10	Down	Overexpressed in autoimmune diseases; inhibits angiogenesis by binding to CXCR3	[Antonelli et al., 2014; Strieter et al., 2005; Yang and Richmond, 2004]
CXCR7	Down*	Binds to CXCL12 to promote angiogenesis; expressed in primitive erythroid cells in embryos, but not expressed in leukocytes in adults	[Berahovich et al., 2010; Strieter et al., 2005]
DSG3	Up	Adhesion molecule in desmosomes; regulates cell-cell adhesion and actin cytoskeleton organization; overexpressed in some cancers; silencing DSG3 slows tumor growth;	[Chen et al., 2013; Chen et al., 2007; Man Tsang et al., 2012; Sumigray et al., 2014; Zhurinsky et al., 2000]
EDN1	Down	Growth factor that activates numerous signaling pathways, including MAPK and EGFR pathways; regulates genes associated with adhesion, angiogenesis, and apoptosis;	[Bagnato et al., 2008; Stow et al., 2011]
EGF	Down	Regulates gene expression in breast cancer cells by binding to EGFR and ERBB2 receptors; causes actin polymerization and membrane ruffling by activating RHO GTPases	[Barry, 2005; Boonstra, 1995; Frušić-Zlotkin et al., 2006; Zhurinsky et al., 2000]

Table 2 continued

Gene	Change in RTA+ Cells	Effect	References
ERBB2	Up**	Overexpressed in some breast cancers; overexpression is associated with increased aggressiveness; regulates gene expression by phosphorylating specific transcription factors in response to EGF	[Barry, 2005]
LAMC2	Up	Extracellular matrix component; typically expressed in epithelial cells	[Colognato and Yurchenco, 2000]
MAP2K1	Down	Increased activation causes dedifferentiation in MDCK cells; inhibition of MAP2K1 in Ras-transformed MDCK cells restores the lost epithelial morphology	[Chen et al., 2000; Schramek et al., 1997]
MMP9	Down	Degrades extracellular matrix proteins; required for the release of vascular endothelial growth factor (VEGF) during long bone development; MMP9-null mice show reduced angiogenesis; increases tumor invasiveness and angiogenesis; inhibited by TIMP's	[Baker, 2002; Chen et al., 2013; Coussens, 2002; McCawley and Matrisian, 2001; Nagase et al., 2006; Patterson and Sang, 1997; Seo et al., 2003; Visse and Nagase, 2003; Yu and Stamenkovic, 2000]
MUC1	Down	Overexpression blocks adhesion to the extracellular matrix; forms part of the mucosal barrier to viral and bacterial infections; interacts with ERBB2 during mammary gland development	[Gendler, 2001; Kumar et al., 2014; Ligtenberg et al., 1992; McAuley et al., 2007; Wesseling et al., 1995]
PDPN	Up	Lymphatic endothelial cell marker; promotes tumor invasion by remodeling the actin cytoskeleton; expression is induced by PROX1 transcription factor	[Hong et al., 2002; Pan et al., 2014; Petrova et al., 2002; Rudzińska et al., 2014; Shindo et al., 2014; Wakisaka et al., 2015; Weninger et al., 1999; Wicki et al., 2006]
PLAU	Up	Involved in suppressing regulatory T cells; produced by endothelial cells during angiogenesis; inhibition of PLAU disrupts metastasis in chick embryos; increased expression correlates with poor prognosis in numerous types of cancer	[Balcerzyk et al., 2011; He et al., 2012; Ulisse et al., 2009]
PRKCD	Up*	Activates the transcription factor estrogen-related receptor α (ERRα) in response to EGF signaling	[Barry, 2005]

Table 2 continued

Gene	Change in RTA+ Cells	Effect	References
TGFB1	Down**	Induces epithelial to mesenchymal transition in hepatocytes; promotes angiogenesis and tumor invasiveness following activation by MMP9; promotes migration, invasion, and filopodia formation in spindle-shaped tumor cells by increasing fascin expression; involved in differentiation and vascular remodeling	[Kojima et al., 2007; Sun et al., 2011; Wang et al., 2004; Yu and Stamenkovic, 2000]
TIMP1 & TIMP2	Up	Prevent MMP's from degrading the extracellular matrix; inhibit angiogenesis independent of MMP blocking actions	[Baker, 2002; Guedez et al., 1998; Nagase et al., 2006; Seo et al., 2003; Visse and Nagase, 2003]
TNFα	(Up)	Proinflammatory cytokine; involved in cancer-related anemia; promotes apoptosis & proliferation	[Kalyani and Jamil, 2015; Lu et al., 2014; Siggs et al., 2012]
UTRN	Up*	Involved in B-cell adhesion to the extracellular matrix; knockdown inhibits adhesion	[Belanto et al., 2014; Costantini et al., 2009]

That screening identified 10 genes clustered in three categories that were potentially biologically significant. Of these 10 genes, 7 were associated with ECM creation and destruction; 1 was associated with cellular adhesion to the ECM; and 2 were associated with epithelial or endothelial differentiation. The RNA-Seq results for the 10 genes were then verified by qPCR (Fig 7).

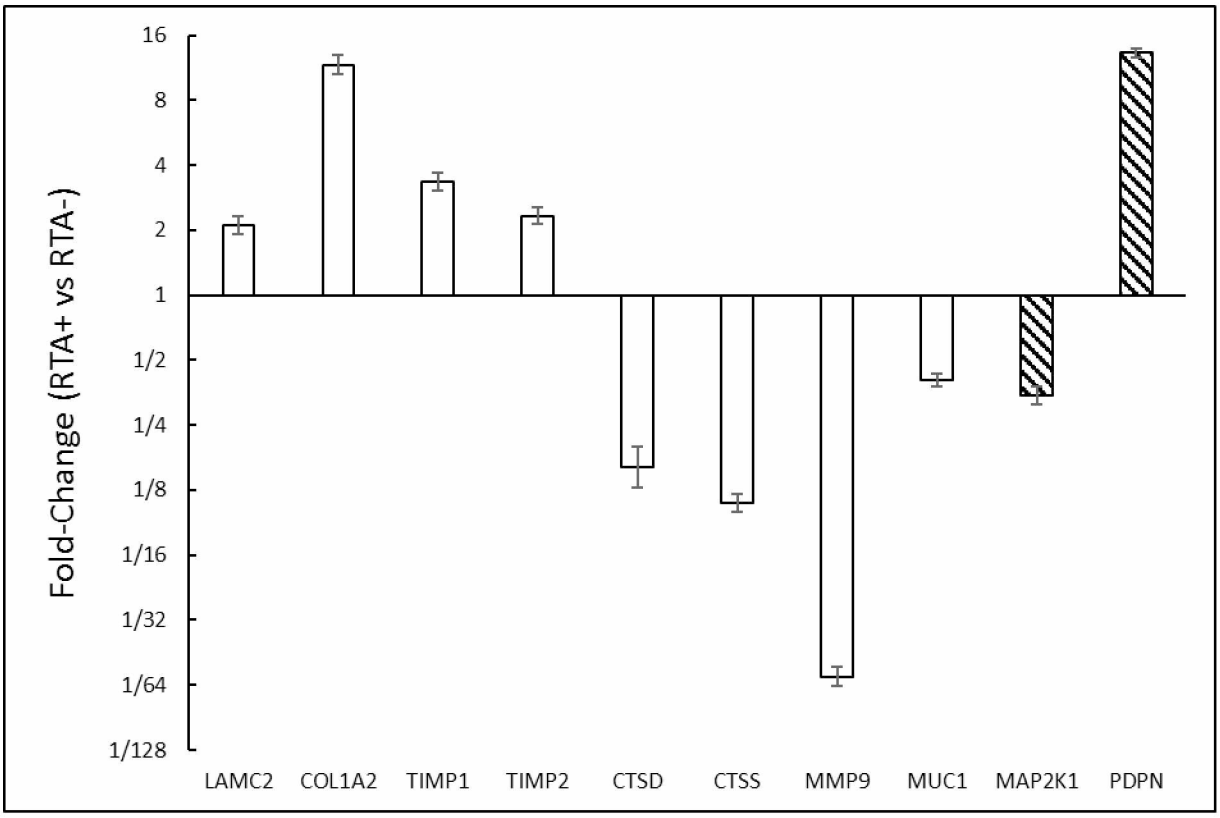

Fig 7: Gene expression fold-changes for potentially biologically significant genes. Differences are between RTA+ and RTA- cell lines as measured by qPCR for genes identified as potentially biologically significant. Error bars represent propagated standard errors. White bars indicate genes involved in ECM remodeling. Striped bars indicate cells involved in epithelial/endothelial differentiation. RTA suppressed genes that degrade or block extra cellular matrix (MUC1, CTSD, CTSS, & MMP9). At the same time, it also increased expression of extra cellular matrix maintenance factors; including both extra cellular matrix components (COL1A2 & LAMC2) and degradation inhibitors (TIMP1 & TIMP2). Additionally, a gene involved in epithelial dedifferentiation (MAP2K1) was down-regulated; while podoplanin (PDPN), a cell marker seen only in lymphatic endothelium, was up-regulated.

2.4.4. ECM Regulation

Three complimentary pathways related to ECM regulation were altered by RTA. ECM degradation genes were down-regulated; genes that block ECM degradation were up-regulated; and ECM components themselves were up-regulated (Fig. 7).

2.4.4.1. Suppression of ECM Degradation Genes

Three genes whose products are involved in degrading the ECM were significantly down-regulated by RTA. Matrix Metallopeptidase 9 (MMP9) was down-regulated over 50-fold in the RTA+ cells. MMPs are a family of zinc-dependent proteases instrumental in the breakdown of the ECM, as well as playing roles in a variety of other cellular functions [Nagase et al., 2006]. Additionally, cathepsins D (CTSD) and S (CTSS) were down-regulated 2.4- and 15-fold respectively. Both cathepsins degrade ECM components. CTSD breaks down ECM proteins collagen and fibronectin releasing the ECM-bound cytokine basic fibroblast growth factor (bFGF) which promotes further degradation of the ECM [Briozzo et al., 1991]. Suppression of CTSS decreases tumor invasiveness [Gocheva et al., 2006] and angiogenesis [Wang et al., 2005].

2.4.4.2. ECM Degradation Blocked by Tissue Inhibitors of Metalloproteases (TIMPs)

In addition to suppressing transcription of genes involved in breaking down the ECM, RTA's presence within MDCK cells stimulated transcription of proteins that inhibit ECM degradation. Tissue inhibitors of metalloproteases (TIMPs) are one such family of proteins. Expression of TIMP1 and TIMP2 was increased 5.4-fold and 4.2-fold respectively. Both have the ability to bind to and inhibit all MMPs, thereby blocking their degradation of the ECM components [Baker, 2002]. The increase in TIMP expression reinforces the decrease in MMP9 expression to further suppress ECM degradation in RTA+ cells.

2.4.4.3. Increased Expression of ECM Components

As well as inhibiting the breakdown of ECM already present, RTA also increased the expression of several genes coding for ECM structural proteins themselves. Expression of the LAMC2 component of laminin was increased 3.6-fold, and expression of the type 1 collagen component COL1A2 was increased 8.7-fold.

2.4.4.4. Cell-Cell and Cell-ECM Adhesion

Mucin 1 (aka Episialin, or MUC1) expression was 78-fold lower in RTA+ cells. MUC1 is a transmembrane protein frequently expressed by epithelial cells as well as a number of hematopoietic cells [Gendler, 2001], and is overexpressed in many carcinomas [Wesseling et al., 1995]. Overexpression of MUC1 has been shown to block both cell-cell and cell-matrix adhesion [Ligtenberg et al., 1992; Wesseling et al., 1995]. MUC1 also forms part of the mucus layers that contribute to the prevention of microbial and viral infection [McAuley et al., 2007].

2.4.5. Differentiation of Epithelial/Endothelial Cells

MAP2K1 (mitogen activated protein kinase kinase 1) is involved in epithelial differentiation. Increased MAP2K1 activity leads to epithelial dedifferentiation in MDCK cells [Schramek et al., 1997] and MAP2K1 was found to be down-regulated 2.4-fold by RTA. Suppression of MAP2K1 was further shown to reverse this dedifferentiation and resulted in restoration of the lost epithelial morphology [Chen et al., 2000]. RTA's ability to suppress MAP2K1 is therefore likely to play a role in MDCK reprogramming.

Podoplanin (PDPN) is a lymphatic endothelial specific marker whose presence in spindle cells was used as evidence that spindle cells originate from lymphatic endothelial cells [Weninger et al., 1999]. Podoplanin was up-regulated 13-fold by RTA.

2.5. Discussion

In this study we identified potential mechanisms by which KSHV was able to induce a spindle-like morphology in host cells. Previous work in our lab had shown that MDCK cells transfected with a plasmid vector coding for KSHV's lytic switch protein (RTA) were reprogrammed to display an elongated shape similar to the spindle cells that predominate in KS tumors. RTA is a transcriptional transactivator and hence the mechanism by which it effects the change in cell shape is likely by modulating gene expression in its host cell. Of the 180 genes up- or down-regulated by RTA, 10 were identified as potentially biologically significant due to their associations with ECM binding and remodeling and/or their identity as BEC- or LEC-specific

markers. These findings support our hypothesis that RTA was effecting the morphological reprogramming of host cells by altering the levels of transcription in signaling pathways associated with regulating cell shape. However, our results can only support a correlation between changes in expression of those 10 genes and the observed changes in cell morphology. Establishing a causative relationship will require selectively targeting each of those genes to counteract the influence of RTA and observing the resulting effect on phenotype.

Not only does RTA modulate expression of genes involved in regulating cell shape (as we predicted), but it also increased expression of the lymphatic endothelial specific protein podoplanin and the blood endothelial specific protein laminin. These results concur with the finding that KSHV infection of either lymphatic endothelial cells (LECs) or blood endothelial cells (BECs) causes those cells to be reprogrammed to display an intermediate gene expression profile that is ~70% LEC-specific and ~30% BEC-specific [Wang et al., 2004]. One limitation of the previous studies demonstrating that KSHV has the ability to reprogram host cells to express lymphatic endothelial specific proteins was that they involved infecting cell cultures with KSHV, rather than transfecting them with a single viral protein as in our study. The effect of that difference was that the presence of the entire virus in the host cells led to a much larger number of gene expression changes; thereby making it much more difficult to identify causal relationships. For example, expression of the master lymphatic vascular development gene PROX1 in blood endothelial cells (BECs) causes those cells to express podoplanin as well as a number of other lymphatic-specific proteins [Hong et al., 2002]. However, when previous whole-virus studies found that infecting BECs with KSHV caused increased expression of both PROX1 and podoplanin, they found that knocking down PROX1 did not fully reduce podoplanin transcription levels back to their initial pre-KSHV levels; indicating that additional factors were driving the increase in podoplanin. Our results indicate that RTA likely causes at least a portion of the remaining unexplained increase in podoplanin expression. Determining whether RTA is the sole remaining factor would require knocking down expression of both PROX1 and RTA in KSHV infected cells and measuring the resulting levels of podoplanin transcription in those cells relative to KSHV-naïve cells.

In this study we demonstrated that RTA is capable of reprogramming host cells to adopt a spindle-like morphology, as well as inducing expression of LEC and BEC specific proteins. Taken together, these findings open the possibility of a non-endothelial origin for KS spindle cells. The current consensus for an endothelial (rather than B-cell) origin for spindle cells is based on morphological similarities between spindle cells and endothelial cells. In order for spindle cells to originate from KSHV-infected B-cells, the virus would have to reprogram the cells to display a spindle cell-like phenotype. Previous work in our lab showed that when PEL derived lymphoblastic BCBL-1 cells were treated with phorbol ester to induce lytic replication, a portion of those cells adopted a spindle cell-like phenotype; and that ectopic expression of RTA in KSHV-naïve cells caused those cells to adopt a spindle cell-like phenotype (J. Chen, unpublished work, 2014). The results of this study further support those findings by identifying a set of gene expression changes induced by RTA that may be causing that phenotypical reprogramming. Further work is needed to determine whether the effects found *in vitro* also exist *in vivo*. If so, then it is possible that, despite their distinctly different phenotypes, KS and PEL may share a B-cell origin.

2.6. References

Abratt RP, Vorobiof DA. 2003. Cancer in Africa. The Lancet Oncology 4:394–396.

Aijaz S, D'Atri F, Citi S, Balda MS, Matter K. 2005. Binding of GEF-H1 to the Tight Junction-Associated Adaptor Cingulin Results in Inhibition of Rho Signaling and G1/S Phase Transition. Developmental Cell 8:777–786.

Anderson MA, Huang D, Roberts A. 2014. Targeting BCL2 for the Treatment of Lymphoid Malignancies. Seminars in Hematology 51:219–227.

Antonelli A, Ferrari SM, Giuggioli D, Ferrannini E, Ferri C, Fallahi P. 2014. Chemokine (C–X–C motif) ligand (CXCL)10 in autoimmune diseases. Autoimmunity Reviews 13:272–280.

Bagnato A, Spinella F, Rosanò L. 2008. The endothelin axis in cancer: the promise and the challenges of molecularly targeted therapyThis article is one of a selection of papers published in the special issue (part 2 of 2) on Forefronts in Endothelin. Canadian Journal of Physiology and Pharmacology 86:473–484.

Baker AH. 2002. Metalloproteinase inhibitors: biological actions and therapeutic opportunities. Journal of Cell Science 115:3719–3727.

Balcerzyk A, Żak I, Emich-Widera E, Kopyta I, Iwanicki T, Pilarska E, Pienczk-Ręcławowicz K, Kaciński M, Wendorff J, Połatyńska K. 2011. The Plasminogen Activator Inhibitor-1 Gene Polymorphism in Determining the Risk of Pediatric Ischemic Stroke – Case Control and Family-Based Study. Neuropediatrics 42:67–70.

Barry JB. 2005. Epidermal Growth Factor-Induced Signaling in Breast Cancer Cells Results in Selective Target Gene Activation by Orphan Nuclear Receptor Estrogen-Related Receptor. Cancer Research 65:6120–6129.

Belanto JJ, Mader TL, Eckhoff MD, Strandjord DM, Banks GB, Gardner MK, Lowe DA, Ervasti JM. 2014. Microtubule binding distinguishes dystrophin from utrophin. Proceedings of the National Academy of Sciences 111:5723–5728.

Benes P, Vetvicka V, Fusek M. 2008. Cathepsin D—Many functions of one aspartic protease. Critical Reviews in Oncology/Hematology 68:12–28.

Berahovich RD, Zabel BA, Penfold MET, Lewen S, Wang Y, Miao Z, Gan L, Pereda J, Dias J, Slukvin II, McGrath KE, Jaen JC, Schall TJ. 2010. CXCR7 Protein Is Not Expressed on Human or Mouse Leukocytes. The Journal of Immunology 185:5130–5139.

Bibel M. 2000. Neurotrophins: key regulators of cell fate and cell shape in the vertebrate nervous system. Genes & Development 14:2919–2937.

Boonstra J. 1995. The epidermal growth factor. Cell Biology International 19:413–430.

Briozzo P, Badet J, Capony F, Pieri I, Montcourrier P, Barritault D, Rochefort H. 1991. MCF7 mammary cancer cells respond to bFGF and internalize it following its release from extracellular matrix: A permissive role of cathepsin D. Experimental Cell Research 194:252–259.

Carroll PA, Brazeau E, Lagunoff M. 2004. Kaposi's sarcoma-associated herpesvirus infection of blood endothelial cells induces lymphatic differentiation. Virology 328:7–18.

Caselli E, Fiorentini S, Amici C, Di Luca D, Caruso A, Santoro MG. 2006. Human herpesvirus 8 acute infection of endothelial cells induces monocyte chemoattractant protein 1-dependent capillary-like structure formation: role of the IKK/NF- B pathway. Blood 109:2718–2726.

Cavailles V, Augereau P, Rochefort H. 1993. Cathepsin D gene is controlled by a mixed promoter, and estrogens stimulate only TATA-dependent transcription in breast cancer cells. Proceedings of the National Academy of Sciences 90:203–207.

Chen YH, Lu Q, Schneeberger EE, Goodenough DA. 2000. Restoration of Tight Junction Structure and Barrier Function by Down-Regulation of the Mitogen-activated Protein Kinase Pathway in Ras-transformed Madin-Darby Canine Kidney Cells. Molecular Biology of the Cell 11:849–862.

Chen Y-J, Lee L-Y, Chao Y-K, Chang JT, Lu Y-C, Li H-F, Chiu C-C, Li Y-C, Li Y-L, Chiou J-F, Cheng A-J. 2013. DSG3 Facilitates Cancer Cell Growth and Invasion through the DSG3-Plakoglobin-TCF/LEF-Myc/Cyclin D1/MMP Signaling Pathway. PLoS ONE 8:e64088.

Chen Y-J, Chang JT, Lee L, Wang H-M, Liao C-T, Chiu C-C, Chen P-J, Cheng A-J. 2007. DSG3 is overexpressed in head neck cancer and is a potential molecular target for inhibition of oncogenesis. Oncogene 26:467–476.

Citi S, Sabanay H, Kendrick-Jones J, Geiger B. 1989. Cingulin: characterization and localization. Journal of Cell Science 93 (Pt 1):107–22.

Citi S, Paschoud S, Pulimeno P, Timolati F, De Robertis F, Jond L, Guillemot L. 2009. The Tight Junction Protein Cingulin Regulates Gene Expression and RhoA Signaling. Annals of the New York Academy of Sciences 1165:88–98.

Citi S, Pulimeno P, Paschoud S. 2012. Cingulin, paracingulin, and PLEKHA7: signaling and cytoskeletal adaptors at the apical junctional complex. Annals of the New York Academy of Sciences 1257:125–132.

Colognato H, Yurchenco PD. 2000. Form and function: The laminin family of heterotrimers. Developmental Dynamics 218:213–234.

Conti I, Rollins BJ. 2004. CCL2 (monocyte chemoattractant protein-1) and cancer. Seminars in Cancer Biology 14:149–154.

Costantini JL, Cheung SMS, Hou S, Li H, Kung SK, Johnston JB, Wilkins JA, Gibson SB, Marshall AJ. 2009. TAPP2 links phosphoinositide 3-kinase signaling to B-cell adhesion through interaction with the cytoskeletal protein utrophin: expression of a novel cell adhesion-promoting complex in B-cell leukemia. Blood 114:4703–4712.

Coussens LM. 2002. Matrix Metalloproteinase Inhibitors and Cancer--Trials and Tribulations. Science 295:2387–2392.

Dittmer DP, Damania B. 2013. Kaposi sarcoma associated herpesvirus pathogenesis (KSHV) - An update. Current Opinion in Virology 3:238–244.

Elliott JT, Tona A, Woodward JT, Jones PL, Plant AL. 2003. Thin Films of Collagen Affect Smooth Muscle Cell Morphology. Langmuir 19:1506–1514.

Ensoli B, Stürzl M. 1998. Kaposi's sarcoma: a result of the interplay among inflammatory cytokines, angiogenic factors and viral agents. Cytokine & Growth Factor Reviews 9:63–83.

Frušić-Zlotkin M, Raichenberg D, Wang X, Frušić-Zlotkin M, Raichenberg D, Wang X, David M, Michel B, Milner Y. 2006. Apoptotic mechanism in pemphigus autoimmunoglobulins-induced acantholysis—possible involvement of the EGF receptor. Autoimmunity 39:563–575.

Gendler SJ. 2001. MUC1, the renaissance molecule. Journal of Mammary Gland Biology and Neoplasia 6:339–53.

Gerke V, Moss SE. 2002. Annexins: From Structure to Function. Physiological Reviews 82:331–371.

Gocheva V, Zeng W, Ke D, Klimstra D, Reinheckel T, Peters C, Hanahan D, Joyce JA. 2006. Distinct roles for cysteine cathepsin genes in multistage tumorigenesis. Genes & Development 20:543–56.

Gospodarowicz D, Greenburg G, Birdwell CR. 1978. Determination of cellular shape by the extracellular matrix and its correlation with the control of cellular growth. Cancer Research 38:4155–71.

Greene W, Kuhne K, Ye F, Chen J, Zhou F, Lei X, Gao S-J. 2007. Molecular biology of KSHV in relation to AIDS-associated oncogenesis. Cancer Treatment and Research 133:69–127.

Guedez L, Courtemanch L, Stetler-Stevenson M. 1998. Tissue inhibitor of metalloproteinase (TIMP)-1 induces differentiation and an antiapoptotic phenotype in germinal center B cells. Blood 92:1342–9.

Guilak F, Cohen DM, Estes BT, Gimble JM, Liedtke W, Chen CS. 2009. Control of Stem Cell Fate by Physical Interactions with the Extracellular Matrix. Cell Stem Cell 5:17–26.

Guillemot L, Hammar E, Kaister C, Ritz J, Caille D, Jond L, Bauer C, Meda P, Citi S. 2004. Disruption of the cingulin gene does not prevent tight junction formation but alters gene expression. Journal of Cell Science 117:5245–56.

Guillemot L, Citi S. 2006. Cingulin regulates claudin-2 expression and cell proliferation through the small GTPase RhoA. Molecular Biology of the Cell 17:3569–77.

Guito J, Lukac DM. 2012. KSHV Rta Promoter Specification and Viral Reactivation. Frontiers in Microbiology 3:1–21.

Hayes MJ, Rescher U, Gerke V, Moss SE. 2004. Annexin-Actin Interactions. Traffic 5:571–576.

He F, Chen H, Probst-Kepper M, Geffers R, Eifes S, del Sol A, Schughart K, Zeng A-P, Balling R. 2012. PLAU inferred from a correlation network is critical for suppressor function of regulatory T cells. Molecular Systems Biology 8:624.

Hong Y-K, Harvey N, Noh Y-H, Schacht V, Hirakawa S, Detmar M, Oliver G. 2002. Prox1 is a master control gene in the program specifying lymphatic endothelial cell fate. Developmental Dynamics 225:351–357.

Ingber DE, Prusty D, Sun Z, Betensky H, Wang N. 1995. Cell shape, cytoskeletal mechanics, and cell cycle control in angiogenesis. Journal of Biomechanics 28:1471–84.

Kalyani P, Jamil K. 2015. A study on biochemical facet of anemia in cancers: A strong link between erythropoietin and tumor necrosis factor alpha in anemic cancer patients. Indian Journal of Cancer 52:127.

Kojima T, Takano K, Yamamoto T, Murata M, Son S, Imamura M, Yamaguchi H, Osanai M, Chiba H, Himi T, Sawada N. 2007. Transforming growth factor-β induces epithelial to mesenchymal transition by down-regulation of claudin-1 expression and the fence function in adult rat hepatocytes. Liver International 28:534–545.

Kumar P, Ji J, Thirkill TL, Douglas GC. 2014. MUC1 Is Expressed by Human Skin Fibroblasts and Plays a Role in Cell Adhesion and Migration. BioResearch Open Access 3:45–52.

Laurent-Matha V, Maruani-Herrmann S, Prébois C, Beaujouin M, Glondu M, Noël A, Alvarez-Gonzalez ML, Blacher S, Coopman P, Baghdiguian S, Gilles C, Loncarek J, Freiss G, Vignon F, Liaudet-Coopman E. 2005. Catalytically inactive human cathepsin D triggers fibroblast invasive growth. The Journal of Cell Biology 168:489–499.

Leal MF, Calcagno DQ, Chung J, de Freitas VM, Demachki S, Assumpção PP, Chammas R, Burbano RR, Smith MC. 2015. Deregulated expression of annexin-A2 and galectin-3 is associated with metastasis in gastric cancer patients. Clinical and Experimental Medicine 15:415–420.

Lecuit T, Lenne P-F. 2007. Cell surface mechanics and the control of cell shape, tissue patterns and morphogenesis. Nature Reviews Molecular Cell Biology 8:633–644.

Liaudet-Coopman E, Beaujouin M, Derocq D, Garcia M, Glondu-Lassis M, Laurent-Matha V, Prébois C, Rochefort H, Vignon F. 2006. Cathepsin D: newly discovered functions of a long-standing aspartic protease in cancer and apoptosis. Cancer Letters 237:167–179.

Ligtenberg MJ, Buijs F, Vos HL, Hilkens J. 1992. Suppression of cellular aggregation by high levels of episialin. Cancer Research 52:2318–24.

Lin CQ, Bissell MJ. 1993. Multi-faceted regulation of cell differentiation by extracellular matrix. The FASEB Journal 7 :737–743.

Lu L, Shi W, Deshmukh RR, Long J, Cheng X, Ji W, Zeng G, Chen X, Zhang Y, Dou QP. 2014. Tumor Necrosis Factor-α Sensitizes Breast Cancer Cells to Natural Products with Proteasome-Inhibitory Activity Leading to Apoptosis. PLoS ONE 9:e113783.

Lukashev ME, Werb Z. 1998. ECM signalling: orchestrating cell behaviour and misbehaviour. Trends in Cell Biology 8:437–41.

Man Tsang S, Brown L, Gadmor H, Gammon L, Fortune F, Wheeler A, Wan H. 2012. Desmoglein 3 acting as an upstream regulator of Rho GTPases, Rac-1/Cdc42 in the regulation of actin organisation and dynamics. Experimental Cell Research 318:2269–2283.

Matter K, Balda MS. 2003. Signalling to and from tight junctions. Nature Reviews Molecular Cell Biology 4:225–237.

McAuley JL, Linden SK, Chin WP, King RM, Pennington HL, Gendler SJ, Florin TH, Hill GR, Korolik V, McGuckin MA. 2007. MUC1 cell surface mucin is a critical element of the mucosal barrier to infection. Journal of Clinical Investigation 117:2313–2324.

McCawley LJ, Matrisian LM. 2001. Matrix metalloproteinases: they're not just for matrix anymore! Current Opinion in Cell Biology 13:534–540.

Mersch J, Jackson MA, Park M, Nebgen D, Peterson SK, Singletary C, Arun BK, Litton JK. 2015. Cancers associated with BRCA 1 and BRCA 2 mutations other than breast and ovarian. Cancer 121:269–275.

Moore PS, Chang Y. 2003. Kaposi's Sarcoma–Associated Herpesvirus Immunoevasion and Tumorigenesis: Two Sides of the Same Coin? Annual Review of Microbiology 57:609–639.

Nagase H, Visse R, Murphy G. 2006. Structure and function of matrix metalloproteinases and TIMPs. Cardiovascular Research 69:562–573.

Ohta M, Kitadai Y, Tanaka S, Yoshihara M, Yasui W, Mukaida N, Haruma K, Chayama K. 2002. Monocyte chemoattractant protein-1 expression correlates with macrophage infiltration and tumor vascularity in human esophageal squamous cell carcinomas. International Journal of Cancer 102:220–224.

Orenstein JM, Alkan S, Blauvelt A, Jeang K-T, Weinstein MD, Ganem D, Herndier B. 1997. Visualization of human herpesvirus type 8 in Kaposi's sarcoma by light and transmission electron microscopy. AIDS 11:F35–F45.

Pan Y, Wang W, Yago T. 2014. Transcriptional regulation of podoplanin expression by Prox1 in lymphatic endothelial cells. Microvascular Research 94:96–102.

Patel S, Xiao P. 2013. Primary Effusion Lymphoma. Archives of Pathology & Laboratory Medicine 137:1152–1154.

Patterson BC, Sang QA. 1997. Angiostatin-converting Enzyme Activities of Human Matrilysin (MMP-7) and Gelatinase B/Type IV Collagenase (MMP-9). Journal of Biological Chemistry 272:28823–28825.

Petrova TV, Mäkinen T, Mäkelä TP, Saarela J, Virtanen I, Ferrell RE, Finegold DN, Kerjaschki D, Ylä-Herttuala S, Alitalo K. 2002. Lymphatic endothelial reprogramming of vascular endothelial cells by the Prox-1 homeobox transcription factor. The EMBO Journal 21:4593–9.

Pollard TD, Cooper JA. 2009. Actin, a Central Player in Cell Shape and Movement. Science 326:1208–1212.

Pyakurel P, Pak F, Mwakigonja AR, Kaaya E, Heiden T, Biberfeld P. 2006. Lymphatic and vascular origin of Kaposi's sarcoma spindle cells during tumor development. International Journal of Cancer 119:1262–1267.

Rochefort H, Cavailles V, Augereau P, Capony F, Maudelonde T, Touitou I, Garcia M. 1989. Overexpression and hormonal regulation of pro-cathepsin D in mammary and endometrial cancer. Journal of Steroid Biochemistry 34:177–182.

Rudzińska M, Gaweł D, Sikorska J, Karpińska KM, Kiedrowski M, Stępień T, Marchlewska M, Czarnocka B. 2014. The Role of Podoplanin in the Biology of Differentiated Thyroid Cancers. PLoS ONE 9:e96541.

Schramek H, Feifel E, Healy E, Pollack V. 1997. Constitutively active mutant of the mitogen-activated protein kinase kinase MEK1 induces epithelial dedifferentiation and growth inhibition in madin-darby canine kidney-C7 cells. Journal of Biological Chemistry 272:11426–33.

Seo D-W, Li H, Guedez L, Wingfield PT, Diaz T, Salloum R, Wei B, Stetler-Stevenson WG. 2003. TIMP-2 Mediated Inhibition of Angiogenesis. Cell 114:171–180.

Shi GP, Villadangos JA, Dranoff G, Small C, Gu L, Haley KJ, Riese R, Ploegh HL, Chapman HA. 1999. Cathepsin S required for normal MHC class II peptide loading and germinal center development. Immunity 10:197–206.

Shindo K, Aishima S, Ohuchida K, Fujino M, Mizuuchi Y, Hattori M, Ohtsuka T, Tokunaga S, Mizumoto K, Tanaka M, Oda Y. 2014. Podoplanin expression in the cyst wall correlates with the progression of intraductal papillary mucinous neoplasm. Virchows Archiv 465:265–273.

Shinohara M, Ohyama N, Murata Y, Okazawa H, Ohnishi H, Ishikawa O, Matozaki T. 2006. CD47 regulation of epithelial cell spreading and migration, and its signal transduction. Cancer Science 97:889–895.

Sick E, Jeanne A, Schneider C, Dedieu S, Takeda K, Martiny L. 2012. CD47 update: a multifaceted actor in the tumour microenvironment of potential therapeutic interest. British Journal of Pharmacology 167:1415–1430.

Siggs OM, Xiao N, Wang Y, Shi H, Tomisato W, Li X, Xia Y, Beutler B. 2012. iRhom2 is required for the secretion of mouse TNF. Blood 119:5769–5771.

Staskus KA, Sun R, Miller G, Racz P, Jaslowski A, Metroka C, Brett-Smith H, Haase AT. 1999. Cellular tropism and viral interleukin-6 expression distinguish human herpesvirus 8 involvement in Kaposi's sarcoma, primary effusion lymphoma, and multicentric Castleman's disease. Journal of Virology 73:4181–7.

Stow LR, Jacobs ME, Wingo CS, Cain BD. 2011. Endothelin-1 gene regulation. The FASEB Journal 25:16–28.

Strieter RM, Burdick MD, Gomperts BN, Belperio JA, Keane MP. 2005. CXC chemokines in angiogenesis. Cytokine & Growth Factor Reviews 16:593–609.

Sumigray K, Zhou K, Lechler T. 2014. Cell-Cell Adhesions and Cell Contractility Are Upregulated upon Desmosome Disruption. PLoS ONE 9:e101824.

Sun J, He H, Xiong Y, Lu S, Shen J, Cheng A, Chang W-C, Hou M-F, Lancaster JM, Kim M, Yang S. 2011. Fascin Protein Is Critical for Transforming Growth Factor Protein-induced Invasion and Filopodia Formation in Spindle-shaped Tumor Cells. Journal of Biological Chemistry 286:38865–38875.

Takei Y, Higashira H, Yamamoto T, Hayashi K. 1997. Mitogenic activity toward human breast cancer cell line MCF-7 of two bFGFs purified from sera of breast cancer patients: co-operative role of cathepsin D. Breast Cancer Research and Treatment 43:53–63.

Ulisse S, Baldini E, Sorrenti S, D'Armiento M. 2009. The Urokinase Plasminogen Activator System: A Target for Anti-Cancer Therapy. Current Cancer Drug Targets 9:32–71.

Visse R, Nagase H. 2003. Matrix Metalloproteinases and Tissue Inhibitors of Metalloproteinases: Structure, Function, and Biochemistry. Circulation Research 92:827–839.

Wakisaka N, Yoshida S, Kondo S, Kita M, Sawada-Kitamura S, Endo K, Tsuji A, Nakanish Y, Murono S, Yoshizaki T. 2015. Induction of epithelial-mesenchymal transition and loss of podoplanin expression are associated with progression of lymph node metastases in human papillomavirus-related oropharyngeal carcinoma. Histopathology 66:771–780.

Wang B, Sun J, Kitamoto S, Yang M, Grubb A, Chapman HA, Kalluri R, Shi G-P. 2005. Cathepsin S Controls Angiogenesis and Tumor Growth via Matrix-derived Angiogenic Factors. Journal of Biological Chemistry 281:6020–6029.

Wang D, Park JS, Chu JSF, Krakowski A, Luo K, Chen DJ, Li S. 2004. Proteomic Profiling of Bone Marrow Mesenchymal Stem Cells upon Transforming Growth Factor β1 Stimulation. Journal of Biological Chemistry 279:43725–43734.

Wang H-W, Trotter MWB, Lagos D, Bourboulia D, Henderson S, Mäkinen T, Elliman S, Flanagan AM, Alitalo K, Boshoff C. 2004. Kaposi sarcoma herpesvirus–induced cellular reprogramming contributes to the lymphatic endothelial gene expression in Kaposi sarcoma. Nature Genetics 36:687–693.

Watt FM. 1986. The extracellular matrix and cell shape. Trends in Biochemical Sciences 11:482–485.

Weninger W, Partanen TA, Breiteneder-Geleff S, Mayer C, Kowalski H, Mildner M, Pammer J, Stürzl M, Kerjaschki D, Alitalo K, Tschachler E. 1999. Expression of vascular endothelial growth factor receptor-3 and podoplanin suggests a lymphatic endothelial cell origin of Kaposi's sarcoma tumor cells. Laboratory Investigation 79:243–51.

Wesseling J, van der Valk SW, Vos HL, Sonnenberg A, Hilkens J. 1995. Episialin (MUC1) overexpression inhibits integrin-mediated cell adhesion to extracellular matrix components. The Journal of Cell Biology 129:255–265.

Wicki A, Lehembre F, Wick N, Hantusch B, Kerjaschki D, Christofori G. 2006. Tumor invasion in the absence of epithelial-mesenchymal transition: Podoplanin-mediated remodeling of the actin cytoskeleton. Cancer Cell 9:261–272.

Yang J, Richmond A. 2004. The angiostatic activity of interferon-inducible protein-10/CXCL10 in human melanoma depends on binding to CXCR3 but not to glycosaminoglycan. Molecular Therapy 9:846–55.

Yu Q, Stamenkovic I. 2000. Cell surface-localized matrix metalloproteinase-9 proteolytically activates TGF-beta and promotes tumor invasion and angiogenesis. Genes & Development 14:163–76.

Zhurinsky J, Shtutman M, Ben-Ze'ev A. 2000. Plakoglobin and beta-catenin: protein interactions, regulation and biological roles. Journal of Cell Science 113 (Pt 1:3127–39.

Chapter 3. General Conclusion

The morphological similarities between KS spindle cells and endothelial cells supported the hypothesis that spindle cells originated form KSHV-infected endothelium. However, previous work in our lab found that KSHV was capable of reprogramming lymphoblastic BCBL-1 cells *in vitro*, causing a portion of them to adopt a spindle cell-like phenotype following lytic reactivation; and that ectopic expression of RTA in KSHV-naïve cells also induced a spindle cell-like phenotype. These observations suggested the possibility that spindle cells may originate from B-cells (or other non-endothelial cells) *in vivo*. The present study investigated the potential mechanisms by which RTA (KSHV's lytic switch protein) could be causing that reprogramming. Analysis of the gene expression changes induced by RTA identified 10 genes potentially responsible for the reprogramming.

Morphological reprogramming being a part of KS formation explains a number of previously puzzling aspects of KS biology. For example, the idea that KS is formed when immunosuppression causes the virus in latently infected endothelial cells to reactivate ignores the fact that KSHV is incapable of forming a stable latent infection in those cells [Grundhoff and Ganem, 2004]. The only cells known to permit a stable latent KSHV infection are B-cells. If KS spindle cells form after infected B-cells have been reprogrammed into endothelial-type cells, then it would be evident why the virus is frequently lost from spindle cells *in vitro* and *in vivo*. Since endothelial cells are not conducive to KSHV latency, once the cell is reprogrammed the latent infection may become unstable and be subsequently lost. Morphological reprogramming may also explain why KS is much more prevalent than PEL. If the same lytic reactivation process is responsible for both the malignant transformation and the morphological reprogramming, then PEL would have to be the result of the lytic reactivation process transforming the host cell, but somehow failing to morphologically reprogram it. If both processes share the same trigger, then extraneous conditions would be required to allow one of them to be activated while the other was not.

In addition to their obvious application to KS and KSHV research, the results of this study have a number of other potential applications as well. For example, because of the narrow focus of

this study on the effects caused by ectopic expression of a single viral protein, the list of interesting genes was much smaller than those typically generated by broadly focused studies (such as GWAS based studies). Not only does that simplify the process of determining which gene expression changes are driving the phenotypic changes, it also ensures that those gene expression changes are actually being caused by RTA. That level of detail in defining triggers and mechanisms involved in cellular reprogramming may be useful in tissue engineering settings.

And lastly, KS is unusual in that it is located approximately halfway along the continuum between healthy cells and cancer cells. Like typical cancers, KS proliferates and invades other tissues; but unlike typical cancers, KS is not genetically unstable and it requires external growth factors to proliferate. That makes it uniquely placed to serve as a model for the investigation of the intermediate steps in malignant transformation. There is a significantly long list of both healthy and cancerous cells to choose from for research purposes, but very few partially transformed cells like KS. The lack of partially transformed cells, combined with the fact that *in vivo* transformation requires a number of steps that often occur months or years apart, makes studying the process of oncogenesis very difficult. Therefore, one of the most valuable aspects of KS research is that it can give us insights into the process of malignant transformation which may be applicable to a vast number of other types of cancer.

References – General Conclusion

Grundhoff A, Ganem D. 2004. Inefficient establishment of KSHV latency suggests an additional role for continued lytic replication in Kaposi sarcoma pathogenesis. Journal of Clinical Investigation 113:124–136.

Appendix A

GeneCards – Set Distiller Results

Table A-1: Genes input into GeneCards Set Distiller Software

Genes Used in Set
ABCB1, ACTB, ADCY6, ADRA1B, ANKHD1, ANXA13, ANXA2, ANXA4, AP3S2, APP, APTX, ARG2, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, ATP7B, AVPR2, BCL2L1, BCL2L2, BEX4, BIRC3, BRCA1, CA2, CASP3, CCL17, CCL2, CCL20, CCL24, CCL7, CD40, CD47, CD97, CGN, CLDN2, CLDN3, CLIC2, CLN3, CLNS1A, COL1A2, COX16, CPNE1, CSF2, CTSC, CTSD, CTSS, CXCL10, CYB5R1, CYP4A11, DAG1, DNAJA1, DSG3, EBAG9, EDN1, EFNB2, EGF, EIF4A1, EIF4EBP3, EPB41, ERBB2, ESRRA, F2RL1, FASTK, FGFR1OP2, FLCN, FOSB, FUT10, FXYD2, FZD6, GALK1, GDI1, GDI2, GNA11, GPX8, GSTA4, GSTP1, GUCA2A, GUCY1B3, GUSB, H19, HLA-DMA, HLA-DMB, HLA-DQB1, HLA-DRB1, HMGN2, HSPB8, IL13RA2, KDM5C, KRT14, KRT19, LAMC2, LGALS9, LIFR, LMAN2, MAL, MAP2K1, MAPK14, MCPH1, MGMT, MMP11, MMP9, MSX2, MT1E, MT2A, MUC1, NACA, NFKBIL1, NPC1L1, OTUB1, PDPN, PFDN1, PFDN5, PFDN6, PLA2G15, PLA2G7, PLAU, PNN, PODXL, POFUT1, POLDIP2, PPP1CA, PPP1CB, PRKACA, PRKCD, PRKDC, PRR5, PSMB8, PSMB9, PTGER2, RAB11A, RAB1A, RAB2A, RBM47, RHBDF2, RHPN2, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SETMAR, SLC1A1, SLC35A3, SLC35B2, SLC35C1, SNAI2, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TBX2, TCOF1, TES, TFF1, TGFB1, THADA, TIMP1, TIMP2, TMBIM6, TMEM47, TMEM57, TNFSF12, TXNRD1, UTRN, WFDC2, ZBED5, ZFPM2

Table A-2: GeneCards Set Distiller output showing a ranked listing of patterns identified in the input list

Descriptor	**Attribute type**	**Score**	**P Value**	**Genes sharing this descriptor**
tumors	DISORDER	96	1.00E-16	ABCB1, ACTB, ANXA2, ANXA4, ARG2, ARHGEF2, ATP2A2, ATP7B, AVPR2, BCL2L1, BCL2L2, BIRC3, BRCA1, CA2, CASP3, CCL17, CCL2, CCL20, CCL24, CCL7, CD40, CD47, CD97, CGN, CLDN2, CLDN3, COL1A2, CSF2, CTSC, CTSD, CXCL10, DAG1, DSG3, EBAG9, EFNB2, EGF, EIF4A1, EPB41, ERBB2, ESRRA, F2RL1, FASTK, FLCN, FOSB, GALK1, GDI1, GSTP1, GUCA2A, GUSB, H19, HLA-DMA, HLA-DQB1, HLA-DRB1, HMGN2, IL13RA2, KRT14, KRT19, LAMC2, LGALS9, LIFR, MAL, MAP2K1, MAPK14, MGMT, MMP11, MMP9, MSX2, MT1E, MT2A, MUC1, NACA, PDPN, PLA2G7, PLAU, PNN, PODXL, PPP1CB, PRKCD, PRKDC, PSMB8, PSMB9, PTGER2, SEPP1, SNAI2, SOCS3, ST13, TES, TFF1, TGFB1, TIMP1, TIMP2, TMBIM6, TNFSF12, TXNRD1, WFDC2, ZFPM2

Table A-2 continued

Descriptor	Attribute type	Score	P Value	Genes sharing this descriptor
protein binding	GO_MOLEC_FUNC	93	1.00E-16	ABCB1, ACTB, ADCY6, ANKHD1, ANXA2, APP, APTX, ARHGEF2, ATP1A1, ATP2A2, ATP7B, AVPR2, BCL2L1, BCL2L2, BIRC3, BRCA1, CA2, CASP3, CCL20, CD40, CD47, CD97, CGN, CLIC2, CLN3, CLNS1A, COL1A2, CSF2, CTSC, CTSD, CXCL10, DAG1, DNAJA1, EDN1, EFNB2, EGF, EIF4A1, EIF4EBP3, EPB41, ERBB2, ESRRA, F2RL1, FASTK, FLCN, FZD6, GDI1, GDI2, GSTP1, HSPB8, KRT14, KRT19, MAL, MAP2K1, MAPK14, MCPH1, MMP9, MSX2, MT2A, MUC1, NACA, NFKBIL1, NPC1L1, OTUB1, PFDN5, PLAU, PODXL, PPP1CA, PPP1CB, PRKACA, PRKCD, PRKDC, PSMB8, PSMB9, RAB11A, RAB1A, RPL27A, RPS10, SETMAR, SLC35A3, SNRPB2, SNRPG, SOCS3, ST13, TBX2, TCOF1, TFF1, TGFB1, TIMP1, TIMP2, TNFSF12, TXNRD1, UTRN, ZFPM2
Placenta	EXPRESSION	87	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, CCL2, CD47, CD97, CLDN3, CLN3, CLNS1A, COL1A2, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EFNB2, EIF4A1, ERBB2, FASTK, FUT10, FZD6, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, KRT14, KRT19, LGALS9, LMAN2, MAL, MAP2K1, MMP11, MSX2, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PLAU, PNN, PODXL, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RBM47, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNAI2, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TBX2, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5

Table A-2 continued

Descriptor	Attribute type	Score	P Value	Genes sharing this descriptor
Whole Blood	EXPRESSION	86	1.00E-16	ACTB, ANXA2, ANXA4, APP, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, BIRC3, CASP3, CCL2, CD47, CD97, CLNS1A, CTSC, CTSD, CTSS, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, F2RL1, FASTK, FOSB, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMA, HLA-DMB, HLA-DQB1, HLA-DRB1, HMGN2, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MAPK14, MMP9, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, PRKCD, PSMB8, PSMB9, PTGER2, RAB11A, RAB1A, RAB2A, RBM47, RHBDF2, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TES, TFF1, TGFB1, TIMP1, TIMP2, TMBIM6, TNFSF12, TXNRD1, ZBED5
B Lymphoblasts	EXPRESSION	86	1.00E-16	ACTB, ANKHD1, ANXA2, ANXA4, APTX, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, BIRC3, BRCA1, CD40, CD47, CD97, CLNS1A, CPNE1, CTSC, CTSD, CTSS, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMA, HLA-DMB, HLA-DQB1, HLA-DRB1, HMGN2, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MAPK14, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PFDN6, PNN, POLDIP2, PPP1CA, PPP1CB, PRKACA, PRKCD, PRKDC, PSMB8, PSMB9, RAB11A, RAB1A, RAB2A, RHBDF2, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TCOF1, TES, TFF1, TGFB1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5

Table A-2 continued

Descriptor	Attribute type	Score	P Value	Genes sharing this descriptor
Myeloid	EXPRESSION	84	1.00E-16	ACTB, ANKHD1, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, CCL2, CD47, CD97, CLN3, CLNS1A, CTSC, CTSD, CTSS, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FOSB, FUT10, GDI1, GDI2, GNA11, GSTP1, GUCA2A, GUSB, H19, HLA-DMA, HLA-DMB, HLA-DQB1, HLA-DRB1, HMGN2, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MAPK14, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, PRKCD, PSMB8, PSMB9, PTGER2, RAB11A, RAB1A, RAB2A, RBM47, RHBDF2, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TES, TFF1, TGFB1, TIMP1, TIMP2, TMBIM6, TNFSF12, TXNRD1, ZBED5
Monocytes	EXPRESSION	84	1.00E-16	ACTB, ANKHD1, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, CCL2, CD47, CD97, CLNS1A, CTSC, CTSD, CTSS, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FOSB, FUT10, GDI1, GDI2, GNA11, GSTP1, GUCA2A, GUSB, H19, HLA-DMA, HLA-DMB, HLA-DQB1, HLA-DRB1, HMGN2, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MAPK14, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PLA2G7, PNN, PPP1CA, PPP1CB, PRKACA, PRKCD, PSMB8, PSMB9, PTGER2, RAB11A, RAB1A, RAB2A, RBM47, RHBDF2, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TES, TFF1, TGFB1, TIMP1, TIMP2, TMBIM6, TNFSF12, TXNRD1, ZBED5

Table A-2 continued

Descriptor	Attribute type	Score	P Value	Genes sharing this descriptor
NK Cells	EXPRESSION	84	1.00E-16	ABCB1, ACTB, ANKHD1, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, BIRC3, CCL2, CD47, CD97, CLNS1A, CTSC, CTSD, CTSS, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FOSB, FUT10, GDI1, GDI2, GNA11, GSTP1, GUCA2A, GUSB, H19, HLA-DMA, HLA-DMB, HLA-DRB1, HMGN2, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MAPK14, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, PRKCD, PRR5, PSMB8, PSMB9, PTGER2, RAB11A, RAB1A, RAB2A, RHBDF2, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TES, TFF1, TGFB1, TIMP1, TIMP2, TMBIM6, TXNRD1, UTRN, ZBED5
Fetal Lung	EXPRESSION	83	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, CCL2, CCL20, CD47, CLDN3, CLNS1A, COL1A2, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EDN1, EFNB2, EIF4A1, FASTK, FUT10, FZD6, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, KRT19, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, MUC1, NACA, OTUB1, PFDN1, PFDN5, PNN, PODXL, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RBM47, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TBX2, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, WFDC2, ZBED5
homeostasis/ metabolism phenotype	PHENOTYPE	83	1.00E-16	ABCB1, ACTB, ADCY6, ADRA1B, ANXA2, APP, APTX, ARG2, ATP1A1, ATP2A2, ATP7B, AVPR2, BCL2L1, BCL2L2, BIRC3, BRCA1, CA2, CASP3, CCL2, CD40, CD47, CGN, CLDN2, CLN3, CSF2, CTSC, CTSD, CTSS, CXCL10, CYP4A11, DAG1, EDN1, EFNB2, EPB41, ERBB2, ESRRA, F2RL1, FASTK, FLCN, GALK1, GDI1, GNA11, GSTA4, GUCY1B3, GUSB, HLA-DQB1, HSPB8, IL13RA2, KRT14, LIFR, MAP2K1, MAPK14, MGMT, MMP11, MMP9, MT2A, MUC1, NACA, NPC1L1, PDPN, PLA2G15, PLAU, PNN, PODXL, PRKACA, PRKCD, PRKDC, PTGER2, RHBDF2, SELK, SELT, SEPP1, SLC1A1, SLC35C1, SOCS3, TBX2, TFF1, TGFB1, TIMP1, TIMP2, TMBIM6, UTRN, ZFPM2

Table A-2 continued

Descriptor	Attribute type	Score	P Value	Genes sharing this descriptor
Dentritic Cells	EXPRESSION	82	1.00E-16	ACTB, ANKHD1, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, BIRC3, CCL2, CD47, CD97, CLNS1A, CTSC, CTSD, CTSS, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FOSB, FUT10, GDI1, GDI2, GNA11, GSTP1, GUCA2A, GUSB, H19, HLA-DMA, HLA-DMB, HLA-DQB1, HLA-DRB1, HMGN2, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MAPK14, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, PRKCD, PSMB8, PSMB9, RAB11A, RAB1A, RAB2A, RBM47, RHBDF2, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TES, TFF1, TGFB1, TIMP2, TMBIM6, TNFSF12, TXNRD1, ZBED5
T Cells (CD8+)	EXPRESSION	81	1.00E-16	ABCB1, ACTB, ANKHD1, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, BIRC3, CD47, CD97, CLNS1A, CTSC, CTSD, CTSS, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FOSB, FUT10, GDI1, GDI2, GSTP1, GUCA2A, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MAPK14, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, PRR5, PSMB8, PSMB9, PTGER2, RAB11A, RAB1A, RAB2A, RHBDF2, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TCOF1, TES, TFF1, TGFB1, THADA, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5
B Cells	EXPRESSION	81	1.00E-16	ACTB, ANKHD1, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BCL2L2, BEX4, BIRC3, CCL2, CD40, CD47, CD97, CLNS1A, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FOSB, FUT10, GDI1, GDI2, GNA11, GSTP1, GUCA2A, GUSB, H19, HLA-DMA, HLA-DMB, HLA-DQB1, HLA-DRB1, HMGN2, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MAPK14, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, PRKCD, PSMB8, PSMB9, RAB11A, RAB1A, RAB2A, RHBDF2, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TES, TFF1, TGFB1, TIMP2, TMBIM6, TXNRD1, ZBED5

Table A-2 continued

Descriptor	Attribute type	Score	P Value	Genes sharing this descriptor
Small Intestine	EXPRESSION	80	1.00E-16	ACTB, ANXA13, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, CCL2, CD47, CLDN3, CLNS1A, COL1A2, CTSC, CTSD, CTSS, CYB5R1, DAG1, DNAJA1, EBAG9, EFNB2, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMA, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, KRT19, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, PSMB8, PSMB9, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5
Tonsil	EXPRESSION	80	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, BIRC3, CCL20, CD47, CLNS1A, COL1A2, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, DSG3, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMA, HLA-DMB, HLA-DQB1, HLA-DRB1, HMGN2, HSPB8, KDM5C, KRT14, KRT19, LGALS9, LMAN2, MAL, MAP2K1, MMP9, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, PSMB8, PSMB9, RAB11A, RAB1A, RAB2A, RHBDF2, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5
T Cells (CD4+)	EXPRESSION	80	1.00E-16	ACTB, ANKHD1, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, BIRC3, CD47, CD97, CLNS1A, CTSC, CTSD, CTSS, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FOSB, FUT10, GDI1, GDI2, GSTP1, GUCA2A, GUSB, H19, HLA-DMA, HLA-DMB, HLA-DRB1, HMGN2, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MAPK14, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, PRR5, PSMB8, PSMB9, PTGER2, RAB11A, RAB1A, RAB2A, RHBDF2, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TCOF1, TES, TFF1, TGFB1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5

Table A-2 continued

Descriptor	Attribute type	Score	P Value	Genes sharing this descriptor
Thymus	EXPRESSION	78	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, CCL2, CD47, CD97, CLNS1A, COL1A2, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMA, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, KRT14, KRT19, LGALS9, LMAN2, MAL, MMP9, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, PSMB8, PSMB9, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5
Retina	EXPRESSION	78	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BCL2L2, BEX4, CA2, CCL2, CD47, CLNS1A, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FOSB, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PDPN, PFDN1, PFDN5, PNN, PODXL, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNAI2, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TES, TFF1, TIMP1, TIMP2, TMBIM6, TMEM47, TXNRD1, ZBED5
Smooth Muscle	EXPRESSION	78	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BCL2L2, CCL2, CCL7, CD47, CLNS1A, COL1A2, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EDN1, EIF4A1, FASTK, FUT10, FZD6, GDI1, GDI2, GNA11, GPX8, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PLAU, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SLC35B2, SNAI2, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5

Table A-2 continued

Descriptor	Attribute type	Score	P Value	Genes sharing this descriptor
Salivary Gland	EXPRESSION	78	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, CCL2, CD47, CLNS1A, COL1A2, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, ESRRA, FASTK, FUT10, FXYD2, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMA, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, KRT14, KRT19, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, NFKBIL1, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, WFDC2, ZBED5
Pineal (Day)	EXPRESSION	78	1.00E-16	ACTB, ANXA2, ANXA4, AP3S2, APP, APTX, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BCL2L2, BEX4, CA2, CCL2, CD47, CLNS1A, CTSC, CTSD, CXCL10, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FOSB, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TMEM47, TXNRD1, WFDC2, ZBED5
Pineal (Night)	EXPRESSION	78	1.00E-16	ACTB, ANXA2, ANXA4, APP, APTX, ARG2, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BCL2L2, BEX4, CA2, CCL2, CD47, CLNS1A, COL1A2, CTSC, CTSD, CXCL10, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FOSB, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TMEM47, TXNRD1, ZBED5

Table A-2 continued

Descriptor	Attribute type	Score	P Value	Genes sharing this descriptor
Bronchial Epithelium	EXPRESSION	77	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, CCL2, CCL20, CD47, CLNS1A, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, DSG3, EBAG9, EIF4A1, FASTK, FUT10, FZD6, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, KRT14, KRT19, LAMC2, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PLAU, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5
Endothelial	EXPRESSION	76	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, ATP7B, BCL2L1, BEX4, BRCA1, CCL2, CD47, CLNS1A, CPNE1, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FOSB, FUT10, GDI1, GDI2, GNA11, GSTP1, GUCA2A, GUSB, H19, HLA-DMA, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MAPK14, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, POLDIP2, PPP1CA, PPP1CB, PRKACA, PRKDC, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TES, TFF1, TIMP1, TMBIM6, TXNRD1, ZBED5
mortality/aging	PHENOTYPE	76	1.00E-16	ABCB1, ACTB, ADRA1B, APP, ATP1A1, ATP2A2, ATP7B, AVPR2, BCL2L1, BIRC3, BRCA1, CA2, CASP3, CCL2, CD47, CD97, CLN3, CLNS1A, COL1A2, CSF2, CTSD, CXCL10, DAG1, DSG3, EDN1, EFNB2, ERBB2, F2RL1, FLCN, FZD6, GNA11, GUCY1B3, GUSB, H19, HLA-DMA, HLA-DQB1, HMGN2, HSPB8, KRT14, KRT19, LAMC2, LIFR, MAP2K1, MAPK14, MCPH1, MGMT, MMP9, MSX2, MT2A, NACA, PDPN, PFDN1, PFDN5, PLAU, PNN, PODXL, POFUT1, PPP1CB, PRKACA, PRKCD, PRKDC, PTGER2, RHBDF2, RPL27A, SELK, SEPP1, SLC35C1, SNAI2, SOCS3, TBX2, TCOF1, TGFB1, TIMP1, TXNRD1, UTRN, ZFPM2

Table A-2 continued

Descriptor	Attribute type	Score	P Value	Genes sharing this descriptor
Bone Marrow	EXPRESSION	75	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, CCL2, CD47, CD97, CLNS1A, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMA, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MAPK14, MMP9, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, PRKCD, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5
Pancreas	EXPRESSION	75	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, CCL2, CCL20, CD47, CLDN3, CLNS1A, COL1A2, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EFNB2, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, KRT19, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5
Uterus	EXPRESSION	75	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, CCL2, CD47, CD97, CLNS1A, COL1A2, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, KRT19, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PDPN, PFDN1, PFDN5, PNN, PODXL, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5

Table A-2 continued

Descriptor	Attribute type	Score	P Value	Genes sharing this descriptor
Fetal Brain	EXPRESSION	75	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BCL2L2, BEX4, CCL2, CD47, CLNS1A, CPNE1, CTSC, CTSD, CXCL10, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUCY1B3, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SLC1A1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5
Amygdala	EXPRESSION	75	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BCL2L2, BEX4, CA2, CCL2, CD47, CLNS1A, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUCY1B3, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SLC1A1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TMEM47, TXNRD1, ZBED5
Sup Cervical Ganglion	EXPRESSION	75	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, CCL17, CCL2, CD47, CLNS1A, COL1A2, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, LGALS9, LIFR, LMAN2, MAL, MAP2K1, MSX2, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SETMAR, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5

Table A-2 continued

Descriptor	Attribute type	Score	P Value	Genes sharing this descriptor
Cardiac Myocytes	EXPRESSION	75	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, CCL2, CD47, CLNS1A, COL1A2, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EDN1, EIF4A1, FASTK, FUT10, FZD6, GDI1, GDI2, GNA11, GPX8, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SNAI2, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5
Fetal Thyroid	EXPRESSION	75	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, CCL2, CD47, CLDN3, CLNS1A, COL1A2, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EFNB2, EIF4A1, FASTK, FUT10, FZD6, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, WFDC2, ZBED5
Pancreatic Islet	EXPRESSION	75	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, CCL2, CCL20, CD47, CLNS1A, COL1A2, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, FXYD2, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUCY1B3, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, KRT19, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5

Table A-2 continued

Descriptor	Attribute type	Score	P Value	Genes sharing this descriptor
Testis Germ	EXPRESSION	75	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, CCL2, CD47, CLNS1A, COL1A2, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, IL13RA2, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, UTRN, WFDC2, ZBED5
Testis Intersitial	EXPRESSION	75	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, CCL2, CD47, CLNS1A, COL1A2, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, IL13RA2, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TES, TFF1, TIMP1, TIMP2, TMBIM6, TMEM57, TXNRD1, UTRN, ZBED5
Prefrontal Cortex	EXPRESSION	74	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BCL2L2, BEX4, CA2, CCL2, CD47, CLNS1A, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUCY1B3, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TMEM47, TXNRD1, ZBED5

Table A-2 continued

Descriptor	Attribute type	Score	P Value	Genes sharing this descriptor
Olfactory Bulb	EXPRESSION	74	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BCL2L2, BEX4, CA2, CCL2, CD47, CLNS1A, COL1A2, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, KRT14, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5
Hypothalamus	EXPRESSION	74	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BCL2L2, BEX4, CA2, CCL2, CD47, CLNS1A, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUCY1B3, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TMEM47, TXNRD1, ZBED5
Fetal Liver	EXPRESSION	74	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, ATP7B, BCL2L1, BEX4, CCL2, CD47, CLNS1A, COL1A2, CTSC, CTSD, CYB5R1, CYP4A11, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, PRKDC, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5

Table A-2 continued

Descriptor	Attribute type	Score	P Value	Genes sharing this descriptor
Trachea	EXPRESSION	74	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, CCL2, CD47, CLNS1A, COL1A2, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, KRT14, KRT19, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, WFDC2, ZBED5
Testis Seminif Tubule	EXPRESSION	74	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, CCL2, CD47, CLNS1A, COL1A2, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, IL13RA2, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, UTRN, ZBED5
Testis Leydig	EXPRESSION	74	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BEX4, CCL2, CD47, CLNS1A, COL1A2, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, IL13RA2, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TBPL1, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, UTRN, ZBED5

Table A-2 continued

Descriptor	Attribute type	Score	P Value	Genes sharing this descriptor
Spinal Cord	EXPRESSION	73	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BCL2L2, BEX4, CA2, CCL2, CD47, CLNS1A, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TMEM47, TXNRD1, ZBED5
Cingulate Cortex	EXPRESSION	73	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BCL2L2, BEX4, CA2, CCL2, CD47, CLNS1A, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUCY1B3, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5
Parietal Lobe	EXPRESSION	73	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BCL2L2, BEX4, CA2, CCL2, CD47, CLNS1A, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUCY1B3, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5

Table A-2 continued

Descriptor	Attribute type	Score	P Value	Genes sharing this descriptor
Temporal Lobe	EXPRESSION	73	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BCL2L2, BEX4, CA2, CCL2, CD47, CLNS1A, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUCY1B3, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5
Occipital Lobe	EXPRESSION	73	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BCL2L2, BEX4, CA2, CCL2, CD47, CLNS1A, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUCY1B3, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5
Cerebellum Peduncles	EXPRESSION	73	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BCL2L2, BEX4, CA2, CCL2, CD47, CLNS1A, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5, ZFPM2

Table A-2 continued

Descriptor	Attribute type	Score	P Value	Genes sharing this descriptor
Thalamus	EXPRESSION	73	1.00E-16	ACTB, ANXA2, ANXA4, APP, ARHGEF2, ATP1A1, ATP2A2, ATP6V0E1, BCL2L1, BCL2L2, BEX4, CA2, CCL2, CD47, CLNS1A, CTSC, CTSD, CYB5R1, DAG1, DNAJA1, EBAG9, EIF4A1, FASTK, FUT10, GDI1, GDI2, GNA11, GSTA4, GSTP1, GUCA2A, GUCY1B3, GUSB, H19, HLA-DMB, HLA-DRB1, HMGN2, HSPB8, KDM5C, LGALS9, LMAN2, MAL, MAP2K1, MT1E, MT2A, NACA, OTUB1, PFDN1, PFDN5, PNN, PPP1CA, PPP1CB, PRKACA, RAB11A, RAB1A, RAB2A, RPL27A, RPS10, RPS15A, RPS24, SELK, SELT, SEPP1, SNRPB2, SNRPG, SOCS3, ST13, TES, TFF1, TIMP1, TIMP2, TMBIM6, TXNRD1, ZBED5

Appendix B

IPA – Networks

Table B-1: Gene networks identified by Ingenuity Pathway Analysis Software

Molecules in Network	**Score**	**Focus Molecules**	**Top Diseases and Functions**
Akt, ANXA2, Ap1, ATP1A1, calpain, CCL2, CCL17, CD3, Cyclin D, ERK1/2, F2RL1, GDI1, GDI2, H19, HLA-DMB, HLA-DR, HLA-DRB1, Ifnar, Interferon alpha, Jnk, MUC1, P38 MAPK, PDGF BB, PDPN, PI3K(complex), Pkc(s), PSMB8, PSMB9, SLC1A1, SLC35B2, Sos, SRC (family), TCR, UTRN, Vegf	43	17	Cell-To-Cell Signaling and Interaction, Hematological System Development and Function, Cell-mediated Immune Response
ANXA4, APP, ATP2A2, BCAR3, BSG, CD300LD, CHTOP, CLNS1A, DDX54, EIF4EBP3, ERK, ERRFI1, ESRRA, FSH, GALK1, GGCT, GSK3B, Ins1, KSR1, LMO4, MAP3K4, Mlcp, NFkB (complex), OTUB1, PLAT, PPP1CB, PPP1R13L, RNA polymerase II, RPS14, RPS25, SMYD3, TFF1, TFIIA, TMBIM6, TNFRSF10B	22	10	Cell Death and Survival, Cellular Compromise, Neurological Disease
AHNAK, ATP1B3, BCS1L, CACNA1C, CGN, CTSB, CYP4A11, DLC1, DNAJA1, EIF4A1, EIF4H, KIAA1524, LPPR4, MED14, MED23, MGMT, MID1, MLH1, NR2C2, OTUD6B, PAIP1, PPAR ligand-PPAR-Retinoic acid-RXRα, PPME1, PTGES, RPL4, RPL7, RPL10, RPS14, RPS19, S100A10, TNFRSF10B, tretinoin, UBASH3B, UBC, UBE3A	7	4	Cancer, Hematological Disease, Lipid Metabolism